8-2013

A NOVEL GAIN OF FUNCTION OF THE *IRX1* AND *IRX2* GENES DISRUPTS AXIS ELONGATION IN THE ARAUCANA RUMPLESS CHICKEN

Nowlan Freese
Clemson University, nowlaxcr4@hotmail.com

A NOVEL GAIN OF FUNCTION OF THE *IRX1* AND *IRX2* GENES DISRUPTS AXIS ELONGATION IN THE ARAUCANA RUMPLESS CHICKEN

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Biological Sciences

---

by
Nowlan Hale Freese
August 2013

Accepted by:
Dr. Susan C. Chapman, Committee Chair
Dr. Lesly A. Temesvari
Dr. Matthew W. Turnbull
Dr. Leigh Anne Clark
Dr. Lisa J. Bain

ABSTRACT

Caudal dysplasia describes a range of developmental disorders that affect normal development of the lumbar spinal column, sacrum and pelvis. An important goal of the congenital malformation field is to identify the genetic mechanisms leading to caudal deformities.

To identify the genetic cause(s) and subsequent molecular mechanisms I turned to an animal model, the rumpless Araucana chicken breed. Araucana fail to form vertebrae beyond the level of the hips. I performed a genome wide association study to identify candidate genomic regions associated with the rumpless phenotype, compared to tailed Araucana. A candidate region of chromosome 2 containing just two genes, *IRX1* and *IRX2*, was identified. In situ hybridization analysis showed that a gain-of-function mutation resulted in both genes being misexpressed at the onset of secondary neurulation in the caudal organizer progenitor population. The caudal progenitor population has a bipotential fate, contributing cells to both mesoderm and neural lineages. This finding is significant because it is the first identified instance of a gain-of-function mutation resulting in axial truncation.

The main question that arises from this novel finding is what is the functional mechanism leading to axial truncation? Possibilities include: the effect on the balance of cell fates within the progenitor population, on proliferation and apoptosis, on cell ingression, and the effect on molecular signaling within caudal tissues. Whereas none of these is

mutually exclusive, I wanted to identify the single molecular event that triggers the cascade of downstream changes that results in axial truncation. I functionally examined each potential to determine the sequence of events in affected Araucana embryos.

Based on the results of this study, I propose a model of development where initial misexpression of the two proneural Iroquois gene family members directs the bipotential progenitor population toward the neural lineage. This results in premature reduction of the progenitor population due to 1) the withdrawal of neuralized cells from the cell cycle, 2) reduced ingression of new progenitor cells via the ventral ectodermal ridge 3) reduced proliferation rates resulting in a failure to extend the axis that then results in 4) early termination of axial elongation and widespread apoptosis.

In conclusion, I have identified a novel genetic basis for axial truncation that sheds light on the molecular mechanisms operating during secondary neurulation and axial elongation.

DEDICATION

I dedicate this to Robert, Betsy, Warren, Caroline, and Daniella Freese. My accomplishments would not have been made, nor would they have carried the same weight, without the support of those I love.

ACKNOWLEDGMENTS

I would like to thank Dr. Susan C. Chapman for not only her help and guidance over the past 5 years, but her belief in my ability to succeed. It is immeasurable how much I have learned, gained and grown over the course of my Ph.D. Thank you.

I would also like to thank everyone on my committee (Dr. Turnbull, Dr. Temesvari, Dr. Bain and Dr. Clark). Each has provided me with a unique perspective on my own work, which has been invaluable over the years.

I am extremely grateful for the help of all of my lab members over the past 5 years. You have made my every day brighter.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

Chapter 3

LIST OF FIGURES

Figure Page

Chapter 1

Chapter 2

List of Figures (Continued)

Figure Page

Chapter 3

CHAPTER ONE

INTRODUCTION

Morphological processes involved in axis elongation

Axis elongation first occurs in embryogenesis during gastrulation. Gastrulation describes the movement of cells through a groove in the epiblast layer of the embryo, called the primitive streak. Epiblast cells ingress through the primitive streak to form the endoderm and mesoderm, forming the triblastic embryo. The remaining epiblast will form the ectoderm layer. Cells of the ectoderm form the neural plate, which will fold to become the neural tube and the surface ectoderm. Following ingression, mesodermal cells will form the head, heart, paraxial, intermediate and lateral plate mesoderm. As the embryo elongates, the presomitic mesoderm becomes segmented into somites in an anterior to posterior direction [1]. Somites are bilateral repeating epithelial spheres that form on either side of the developing neural tube. Once formed, the somites are no longer considered as presomitic mesoderm, rather they become paraxial mesoderm. Somites differentiate to form dermis, muscle, and vertebrae [1,2] **(Fig. 1A)**. Somites can be divided into two primary compartments, the sclerotome (forms the axial skeleton) and dermomyotome (forms the dorsal dermis and skeletal muscles) **(Fig. 1A)** [2,3]. As the numbers of somites can be readily quantified, they provide a visual measure of progress of axis elongation.

In chicken, the first four somites and the anterior half of the fifth somite contribute to the occipital bone, but do not contribute to vertebrae [4]. The remaining somites contribute to

form thirteen cervical vertebrae, seven thoracic vertebrae, fourteen lumbar and sacral vertebrae (lumbar and sacral vertebrae fuse together), five caudal vertebrae and six posterior most vertebrae, which fuse to form the mature pygostyle [5]. Prior to formation of the vertebrae, the sclerotome portion of the somites undergoes resegmentation [6]. The anterior sclerotome of one somite will migrate and fuse with the posterior sclerotome of the somite anterior to it, and this resegmented structure will form the associated vertebra. The posterior sclerotome is cell dense, and does not allow migration of elongating motor axons from the neural tube through, necessitating resegmentation of the anterior and posterior halves of the sclerotome (referred to as von Ebner's fissure) to allow for nervous innervation of the periphery [6].

Somites form at regular intervals during development, with the rate of formation differing between species [7]. For example, zebrafish form one pair of somites every 30 minutes, compared to chicken every 90 minutes and mice every 120 minutes [1,7]. Chickens form between 51-53 somites during embryogenesis. Since somites form the vertebrae, differences in somite number equates to variation in the number of vertebrae between species. For example, some species of frogs having only 6-9 presacral vertebrae, compared to the common corn snake, which has 226 presacral vertebrae **(Fig. 1B-C)** [7,8].

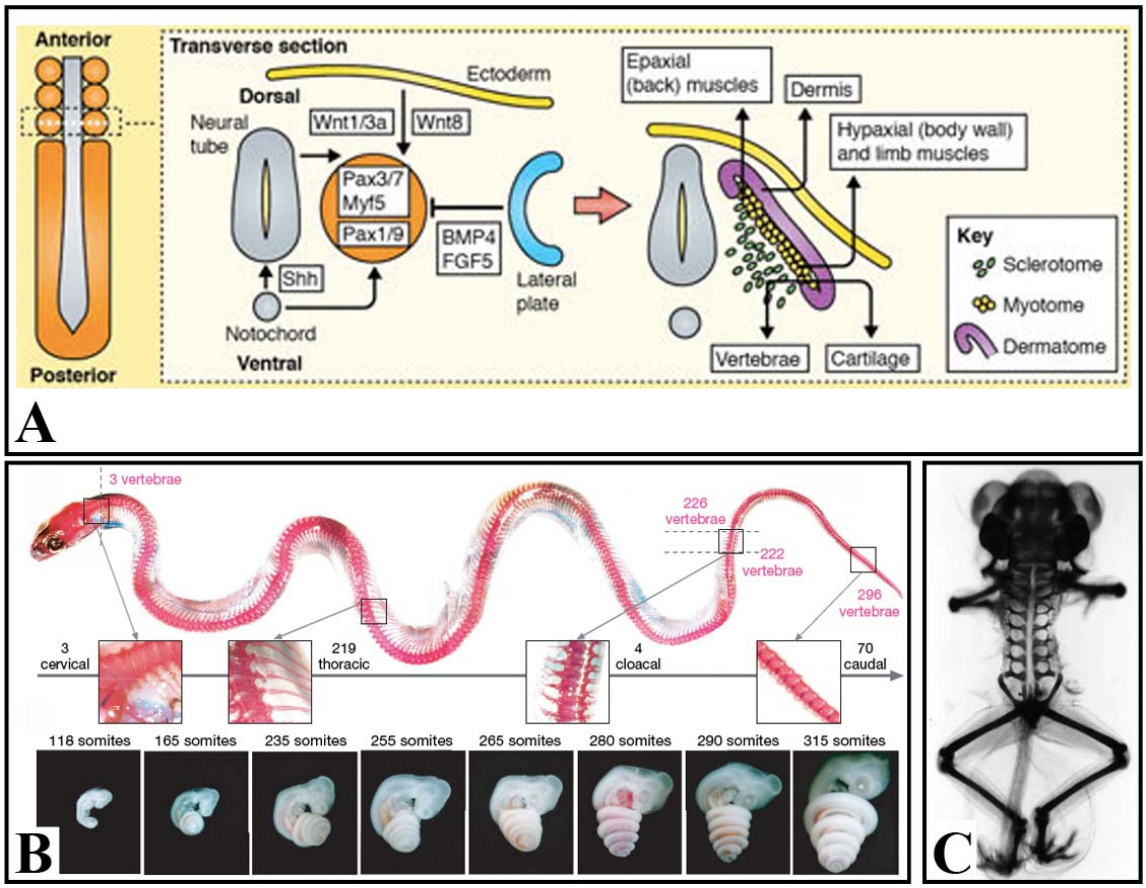


Figure 1: Somites are the precursors to the vertebrae.

A: Somites form from anterior to posterior, segmenting the presomitic mesoderm. Somites break down to form the sclerotome, myotome, and dermatome. The sclerotome forms the vertebrae, the myotome forms muscle, and the dermatome gives rise to the dermis. Adapted from Maroto et al., 2012 [2].

B: The corn snake skeleton showing over 200 vertebrae. The number of somites formed during embryogenesis is associated with the number of vertebrae formed. Adapted from Gomez et al., 2008 [7].

C: Frog skeleton illustrating a low number of vertebrae. Adapted from Richardson et al., 1998 [8].

Morphogenesis of the caudal embryonic axis occurs during posterior axis elongation

The staging of chicken embryos uses the Hamburger and Hamilton (HH) table of normal stages [9]. Images of embryonic stages that I will commonly refer to are shown below (Fig. 2). Primary body formation begins at gastrulation and is followed by the formation of the head, trunk and limbs [1]. Secondary body formation describes the process of body elongation beyond the anus, which requires a distinct process from primary body formation. In chicken, secondary body formation begins when the remnant of the primitive streak has regressed to the posterior of the embryo and forms a mass of mesenchymal cells, which then becomes the tailbud at approximately HH14 (**Fig. 2 and 3A)** [9,10,11]. This early tailbud is defined as a dense posterior group of cells continuous with the most posterior extend of the neural tube and notochord [10]. The tailbud itself then elongates becoming delineated from the surrounding tissue **(Fig. 3B-C)**. Sections of the tail reveal that the neural tube runs posterior until it meets the rostral tailbud, which appears as a mesenchymal cell mass **(Fig. 3D-G)**. In chicken embryos, the tail continues to grow, eventually curving 180° underneath the ventral body to point towards the head. Only later does the tail straighten out in line with the primary body axis. The process of secondary body formation is unique in that both neural cell lineages and mesoderm lineages arise from a single mesenchymal progenitor population.

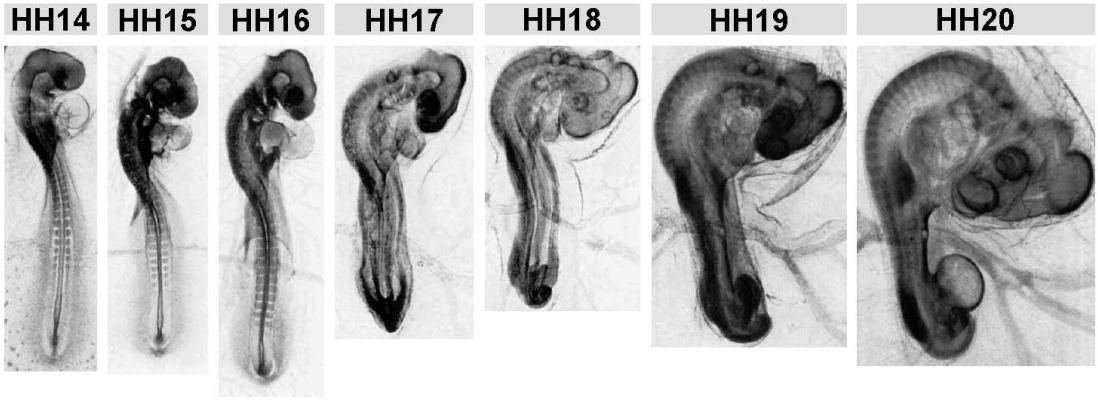


Figure 2: Hamburger and Hamilton (HH) stages of chicken embryo development.

HH14-20. For older embryos, embryonic day (E) will be used as a reference to the number of days the egg has been incubated. Anterior to top. Note the growth of the tail during these stages, which curves ventrally to point back towards the head of the embryo. Image from Hamburger and Hamilton, 1951 [9].

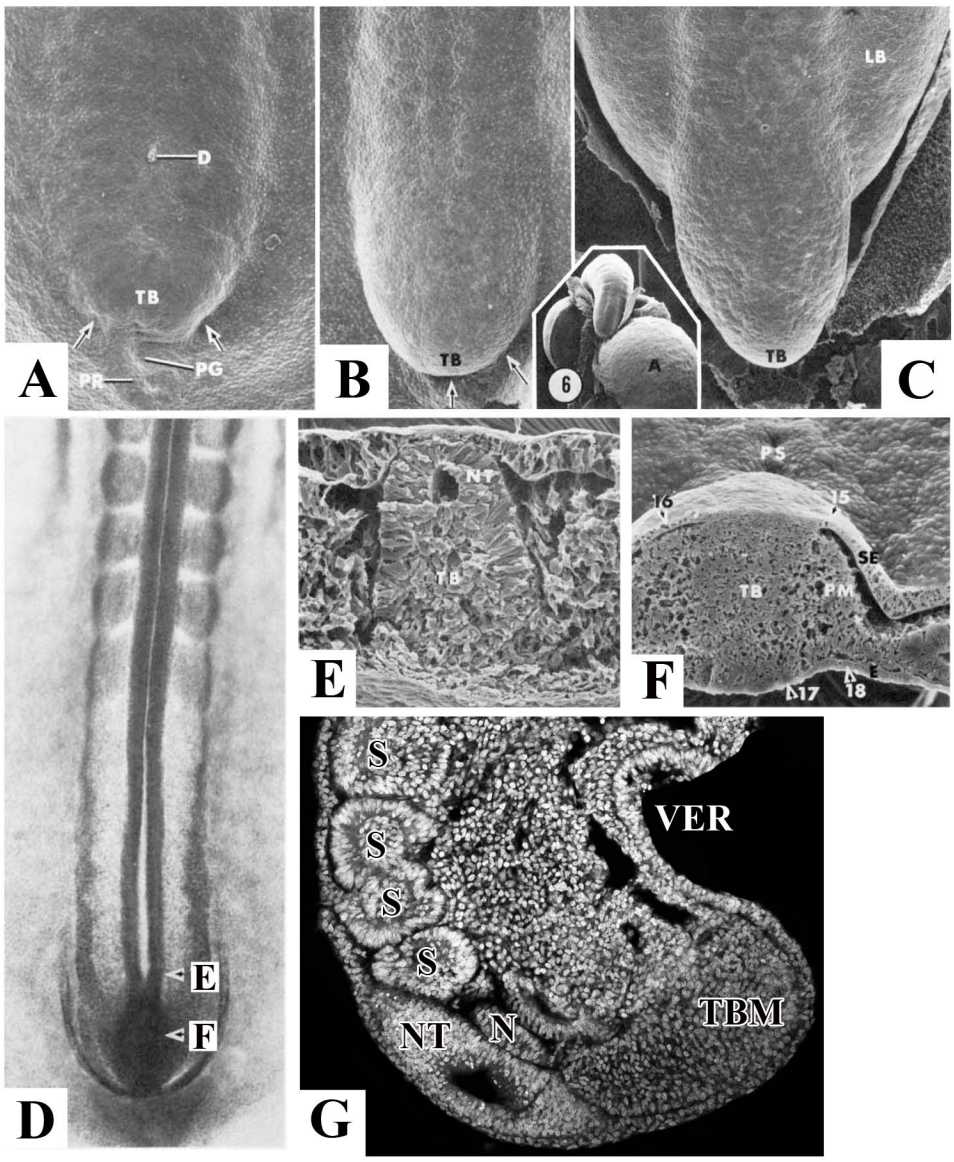

A
D
TB
PR
PG
B
TB
6
A
C
LB
TB
E
NT
TB
F
PS
16
15
SE
TB
PM
17
18
E
D
E
F
G
S
S
S
S
NT
N
VER
TBM

Figure 3: Morphogenesis of the tailbud.

A: Arrows denote the forming posterior body fold of the tailbud at HH14.

B: Delineation of the tailbud from the surrounding blastoderm at HH15.

C: At HH17 tailbud is clearly delineated from surrounding tissue, growth continues, and the limb buds become observable. Insert is image of HH20, ventral view.

D: Dorsal view of HH15 embryo.

E: Transverse section from D. Taken at level of posterior neural tube and the beginning of the tailbud. Note the end of the neural tube, and the beginning of the tailbud mesenchyme.

F: Transverse section from D. Taken at level of tailbud. The tailbud consists of mesenchymal tissue that is beginning to form the paraxial mesoderm and neural tube.

Images A-F adapted from Schoenwolf, 1979 [11].

G: Sagittal section of HH20 tail.

A-D posterior at bottom of image. E-F dorsal up, G dorsal to left. A-allantois, LB-leg bud, N-notochord, NT-neural tube, PG-primitive groove, PM-paraxial mesoderm, PR-primitive ridge, PS-primitive streak, S-somite, SE-surface ectoderm, TBM-tailbud mesenchyme, VER-ventral ectodermal ridge.

Differences between primary versus secondary body formation

During primary body formation the presomitic mesoderm is derived from the gastrulating cells that ingress through the primitive streak and the neural tube is derived from the ectoderm of the neural plate [1]. During secondary body formation, the presomitic mesoderm is derived from a combination of cell ingression through the ventral ectodermal ridge (VER) and proliferation within the tail, specifically the progenitor population [12,13,14]. The VER arises from the remnant of the primitive streak as it reaches the posterior of the embryo and involutes to become the ventral part of the tail **(Fig. 4A)** [15,16]. The VER appears in sections as an ectodermal thickening continuous with the adjacent tail mesenchyme [14].

In chicken embryos, formation of the VER occurs between HH16-18. The ectoderm and tail mesenchyme are separated by a basal lamina. Additionally, the ectoderm is held tightly together by cell-cell adhesion molecules [12]. As the VER forms, cells undergo an epithelial to mesenchymal transition and migrate into the tail mesenchyme **(Fig. 4F-G)** [12]. This requires the breakdown of the basal lamina as well as the down regulation of E-cadherin, a cell-cell adhesion protein **(Fig. 4B-E)**. Fate mapping shows that these cells contribute to the tailbud mesenchyme, paraxial mesoderm, and gut tissue. Migration of cells through the VER and into the tailbud mesenchyme continues until HH24, when the basal membrane is re-established and cells are prevented from migrating through the ventral ectodermal ridge. Ingression of cells through the VER into the tail mesenchyme is considered the final phase of gastrulation.

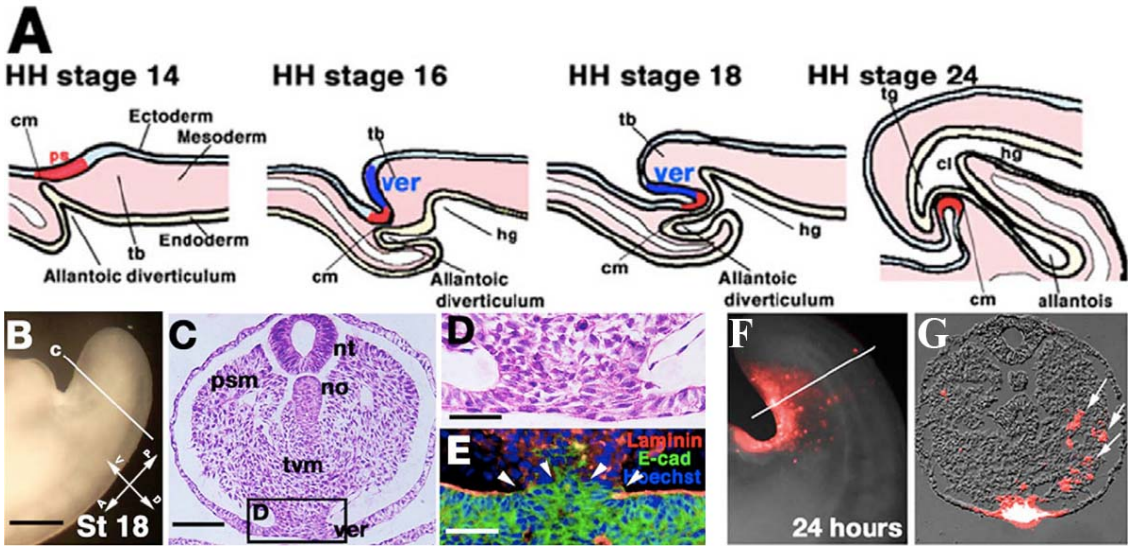


Figure 4: Formation of, and migration into, the ventral ectodermal ridge.

A: Sagittal illustration at HH14 showing how the remaining primitive streak (red) begins to fold under ventrally. During folding, the primitive streak contributes to the forming VER (blue). Anterior to right.

B: Sagittal whole mount HH18 chick tailbud. Scale bar: 400µm..

C: Transverse section of tailbud at level indicated in B. Scale bar: 100µm.

D: Zoomed view of VER. Scale bar: 50µm.

E: Transverse section of the breakdown of laminin and E-cadherin. Scale bar: 50µm.

F: HH20 chick tailbud labeled with DiI.

G: Transverse section at level indicated in F. DiI labeled cells (arrows) can be seen migrating into the tail through the VER. Above images adapted from Ohta et al., 2007 [12].

Abbreviations: cl-cloaca, cm-cloacal membrane, hg-hind gut, no-notochord, nt-neural tube, ps-primitive streak, psm-presomitic mesoderm, tb-tailbud, tvm-tail ventral mesoderm, ver-ventral ectodermal ridge.

During primary body formation the primitive streak is comprised of the primitive pit, the primitive node, and the primitive groove. The node-streak border consists of the primitive node and the anterior primitive groove. Cells from this area contribute to the somites, neural tube, and notochord **(Fig. 5A)** [15,17,18]. In addition, the caudo-lateral epiblast, which extends posterior from the node-streak border, contributes to the neural tube and somitic tissue **(Fig. 5A)** [18]. During secondary body formation, the chordoneural hinge comprises of the posterior ventral neural tube and notochord, and is derived from the node-streak border and the caudo-lateral epiblast **(Fig. 5B-B')** [18]. The chordoneural hinge also contributes cells to the neural tube, notochord, and somites, and is proliferative throughout tailbud development **(Fig. 5C)** [13,15]. As the chordoneural hinge contributes cells to both the mesoderm (somites) and ectoderm (neural tube), this suggests that these cells are multipotent. Further evidence comes from chordoneural hinge derived cells being able to form both neural and mesodermal tissue even after passaging multiple times in cell culture [13,19].

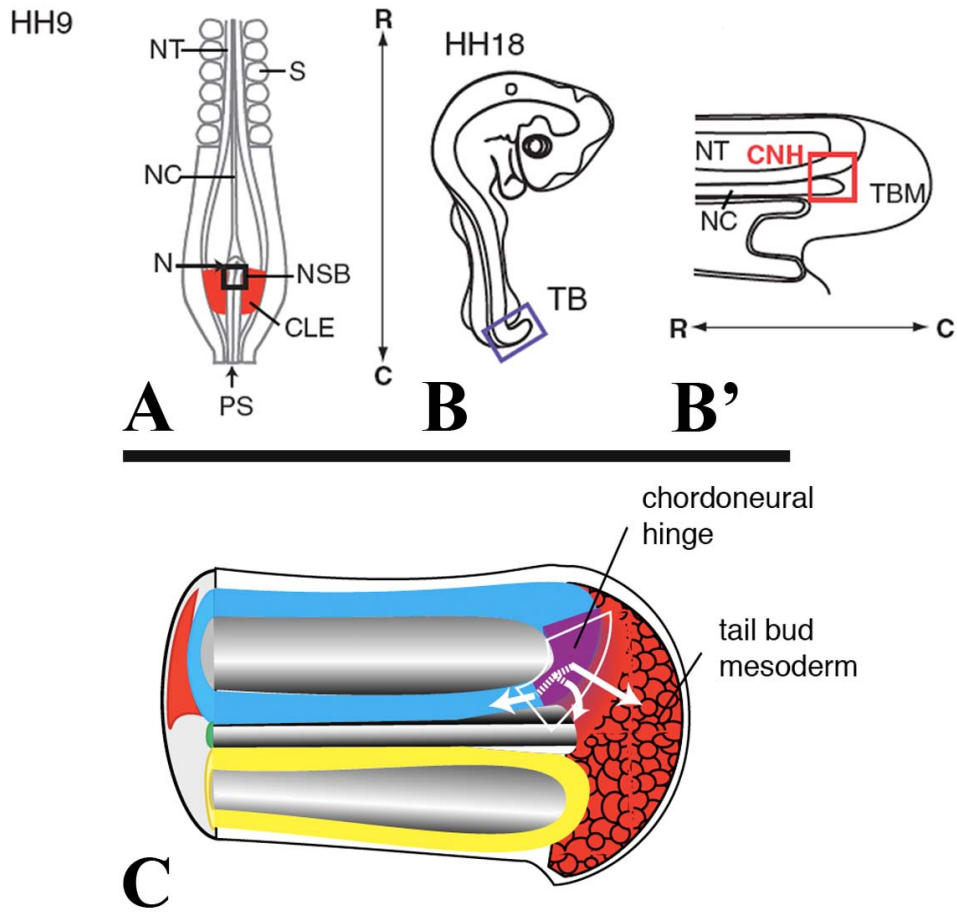


Figure 5: Formation and contribution of the chordoneural hinge.

A: HH9 chick embryo with node-streak border (NSB) and caudal lateral epiblast (CLE). Both of these populations contribute to the future chordoneural hinge (CNH).

B: HH18 chick embryo.

B': View of tail, showing position of CNH (red box). The CNH consists of the caudal posterior neural tube (NT) and the posterior notochord (NC). A-B' adapted from Wilson et al., 2009 [18].

Abbreviations: R-rostral, C-caudal, TB-tailbud, S-somite, PS-primitive streak, TBM-tailbud mesoderm, N-node.

C: Mouse embryo tail showing CNH. The CNH contributes to the neural tube, notochord, and tailbud mesoderm. Image adapted from Cambray and Wilson, 2002 [19]. Anterior to left.

During primary body formation, primary neurulation occurs with the formation of the neural tube from the neural plate [20,21]. Formation of the neural tube occurs when folding of the neural plate first creates the neural groove, which acts as the ventral hinge point around which the neural folds rise up, meeting at the dorsal lips, to form a hollow tube [21]. This differs from formation of the neural tube during secondary body formation (secondary neurulation), where the neural tube forms from the mesenchyme of the tail [16,22]. Cells located dorsally within the tail are specified to become neural, and form into a rod like structure called the medullary cord **(Fig. 6A)**. The medullary cord undergoes cavitation to form an open lumen, which then becomes connected to the lumen of the primary neural tube, forming the complete neural tube [11,16,22].

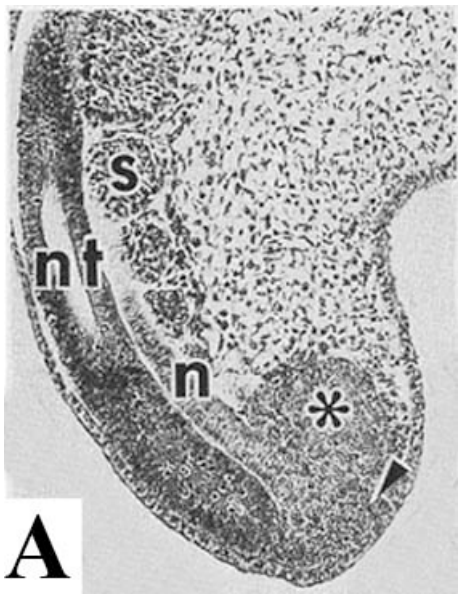


Figure 6: Formation of the medullary cord.

A: Sagittal section of E4 chicken embryo. The posterior of the medullary cord can be seen at the dorsal tip of the tail (arrow). Dorsal to the left, anterior up. Asterisk-tailbud mesenchyme. Abbreviations: nt-neural tube, n-notochord, s-somite. Adapted from Schoenwolf, 1981 [16].

Bipotential fate model

Tissues such as the node-streak border and later chordoneural hinge contribute cells to both a mesoderm (somites) and ectoderm (neural) fate [13,18]. During secondary body formation, posterior FGF signals act to maintain the mesenchyme cells of the tailbud in an undifferentiated state [23,24]. Expression of a number of genes is required to direct these cells to form either mesoderm or neural lineages. Expression of *TBX6*, *WNT3A*, and brachyury is required to maintain mesoderm identity and form the somites, the loss of which causes truncation of the axis [25,26,27,28,29,30,31,32].

Knockdown of *TBX6* leads to a switch in cell fate from mesodermal to neural [29,30]. In *TBX6* knockout mice, instead of forming somites the paraxial mesoderm forms two ectopic neural tubes on either side of the neural tube [30]. In mice, TBX6 acts to inhibit the N1 enhancer of the neural gene, *Sox2,* in cells that will form the paraxial mesoderm, whereas the cells that are not exposed to TBX6 will express *Sox2* and are fated to become neural **(Fig. 7A)** [27].

Knockdown of *WNT3A* expression leads to a decrease in the paraxial mesoderm and an increase in neural tissue in the form of an expanded neural tube **(Fig. 7B)** [28,33]. Similar to the inhibition of *TBX6*, inhibition of *WNT3A* causes cells to contribute to the neural lineage instead of mesoderm [33,34]. However, *WNT3A* inhibition does not lead to ectopic neural tubes as in *TBX6* knockouts. Rather, the neural cells fail to migrate

laterally, and instead form a mass of neural tissue medially expanding the diameter of the neural tube [29,35].

Furthermore, WNT3A is involved in a regulatory loop with brachyury, whereby downstream effectors of WNT3A signaling directly bind the brachyury promoter and prevent transcription [26]. Loss of brachyury leads to a failure to form mesoderm, resulting in axis truncation as can be seen in the T knockout mouse [26,32]. These data strongly suggest that the progenitor population involved in axis elongation consists of cells capable of contributing to either mesodermal or neural tissues.

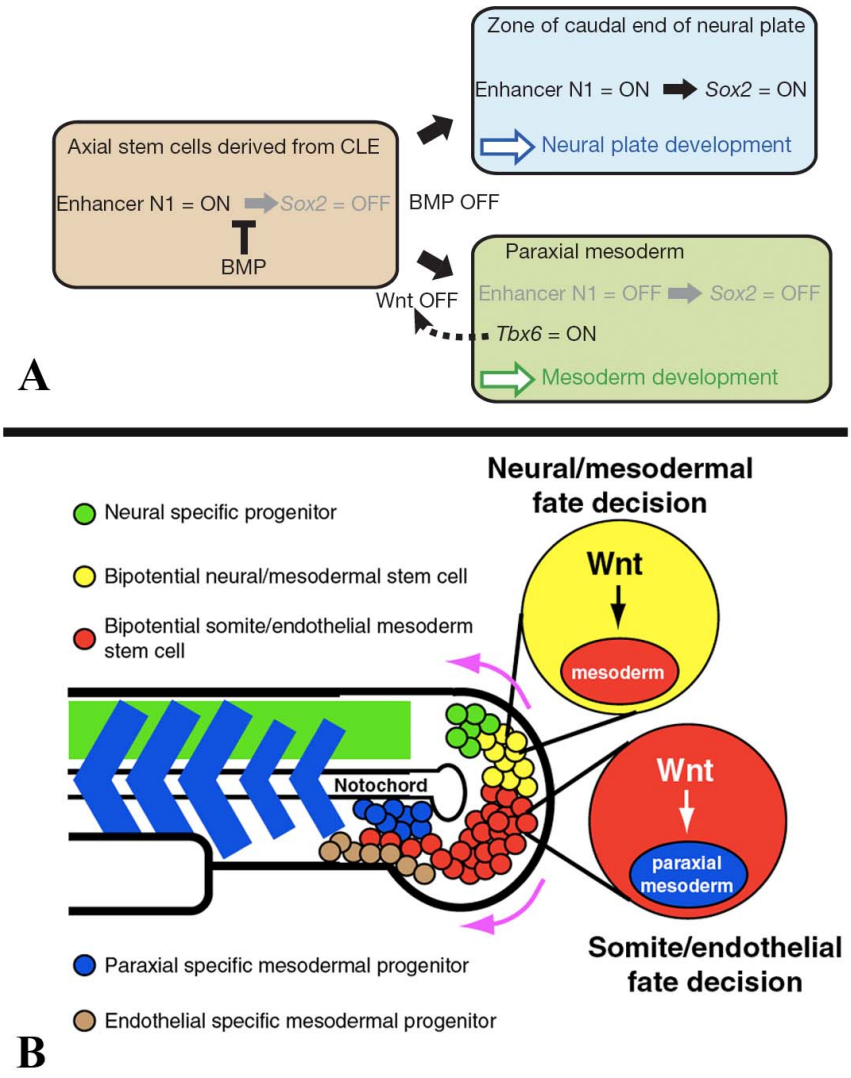


Figure 7: Bipotential fate model.

A: The role of enhancer N1 in the regulation of *SOX2*, and the determination of neural versus mesoderm. Axial stem cells exposed to SOX2 form neural tissue. In the presence of TBX6, enhancer N1 is blocked and *SOX2* is not expressed. This results in mesoderm developing. Adapted from Takemoto et al., 2011 [27].

B: A model of bipotential fate in zebrafish. WNT signaling acts to specify a mesodermal fate in the bipotential cells of the tailbud. WNT signaling further specifies paraxial mesoderm from the pool of mesodermal precursors. Image from Martin and Kimelman, 2012 [33].

Signaling and cell cycling processes involved in axis elongation

The process of axis elongation and somitogenesis are heavily dependent on signaling gradients and the correct expression of transcription factors to maintain a pool of progenitor cells in order to continue elongation, and to properly differentiate those cells that will give rise to either the paraxial mesoderm or the neural tube. Significant work has been done to model and understand the signals and patterning required for proper somitogenesis [1,18,36].

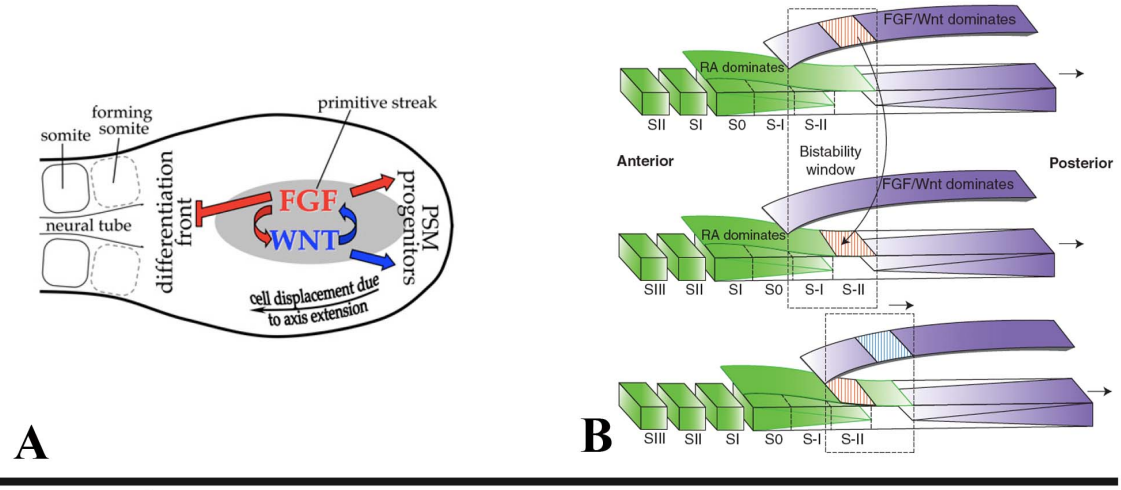
The regular formation of somites every 90 minutes in chicken embryos suggests a mechanism for precisely controlling differentiation and epithelialization of somites. Cells within the tailbud remain undifferentiated and as they move into the presomitic mesoderm they maintain their undifferentiated state [1,36]. However, as the axis continues to extend and these cells and the determination front coincide they begin to differentiate and form somites. This suggests two processes. The first is that a gradient exists to identify cells as being anterior/posterior. The second process requires a regular or rhythmic occurrence that would cause cells to differentiate into somites every 90 minutes. The clock and wavefront model of somitogenesis describes these processes [1,37].

Determination front

The anterior posterior identity of the somite is determined by a dual gradient. FGF, a secreted signaling factor, is highly expressed in the developing tailbud **(Fig. 8A-B)**

[23,24,38]. There, FGF and WNT signaling act to co-regulate each other; with FGF signaling maintaining cells in an undifferentiated state [23]. Two FGF secreted signaling factors set the anterior limit of the determination front; FGF8 and FGF4. These were determined in mice mutants where severe shifts in the wavefront, as well as premature differentiation of the presomitic mesoderm, were only present in mice with a dual knockout of *FGF*4 and *FGF*8 [23,24].

In opposition to the low anterior/high posterior gradient of FGF, a gradient of retinoic acid exists. Retinoic acid is increased anteriorly and lower posteriorly **(Fig. 8B)**. This is due to the expression of *Raldh2*, a retinoic acid-synthesizing enzyme, which is highly expressed in the rostral paraxial mesoderm [39,40]. In addition, *CYP26A1*, a cytochrome p450 enzyme that degrades retinoic acid is highly expressed within the tailbud [41]. Studies in chicken embryos using ectopic FGF protein, or inhibitors of retinoic acid, resulted in a lack of differentiation of somites from the presomitic mesoderm [42]. These opposing gradients form the determination front, the point at which cells become capable of segregating based on their anterior position within the presomitic mesoderm [43]. As the determination front moves caudally the rostral most presomitic cells are now exposed to low levels of FGF and high levels of retinoic acid, resulting in a change in the cells' bistability state [1,36]. At this point the presomitic mesoderm cells are primed to undergo segmentation into somites, but are awaiting an additional signal.


somite
forming somite
primitive streak
neural tube
differentiation front
FGF
WNT
PSM progenitors
cell displacement due to axis extension
A
FGF/Wnt dominates
RA dominates
SII SI S0 S-I S-II
Anterior
Bistability window
Posterior
FGF/Wnt dominates
RA dominates
SIII SII SI S0 S-I S-II
SIII SII SI S0 S-I S-II
B

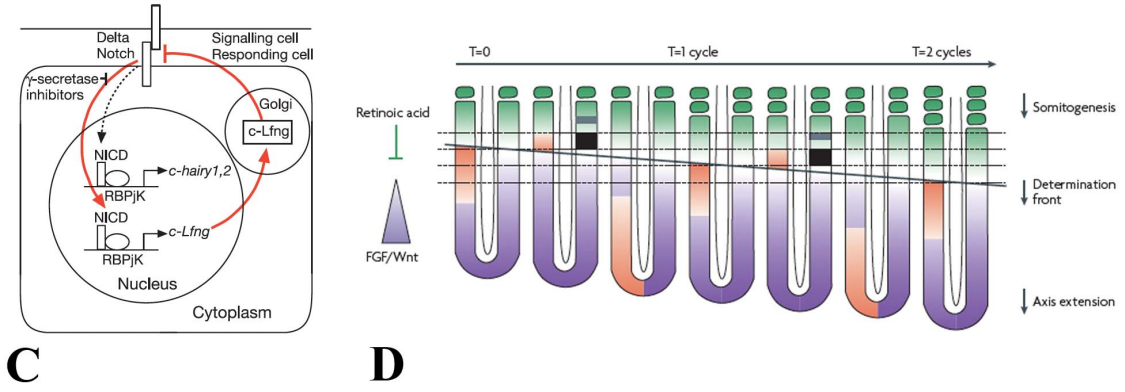

Delta
Notch
Signalling cell
Responding cell
γ-secretase inhibitors
Golgi
c-Lfng
NICD
RBPjK
c-hairy1,2
NICD
RBPjK
c-Lfng
Nucleus
Cytoplasm
C
T=0
T=1 cycle
T=2 cycles
Retinoic acid
FGF/Wnt
Somitogenesis
Determination front
Axis extension
D

Figure 8: Clock and Wavefront model of somitogenesis.

A: FGF and WNT signaling coregulate each other within the tailbud, as well as maintain the PSM. As the axis elongates and cells are displaced more anteriorly, FGF no longer inhibits differentiation and cells begin to form epithelial somites. Image from Naiche et al., 2011 [23].

B: Model for bistability states between FGF/WNT and retinoic acid. As the levels of FGF/WNT decrease and retinoic acid decrease, there is a switch in the state of the cells. A periodic trigger, allowing the sudden change, stimulates this rapid switch. Image from Aulehla and Pourquié, 2010 [36].

C: Loop between *NOTCH* and *LFNG* creates a cyclical loop, or clock. *NOTCH* activates cyclical gene expression, but is then down regulated by up regulated *LFNG*. Adapted from Dale et al., 2003 [44].

D: The clock and wavefront model. Cyclic (clock) expression drives a traveling (wavefront) wave of expression from posterior to anterior. The determination front (black line) is set by the decreasing gradient of FGF and increasing retinoic acid. When the traveling wavefront reaches the determination front, those cells are triggered to differentiate and form epithelial somites. The clock then resets and the process begins again, continuously moving posterior as the axis elongates. Anterior to top. Image from Dequéant and Pourquié, 2008 [1].

Molecular oscillator (clock) is responsible for the timing of somite formation

Only once the determination front has moved caudally, beyond the rostral most presomitic mesoderm cells are they capable of responding to the molecular oscillator, or clock. *Notch1* has been implicated in the cyclic gene network as periodic expression of genes downstream of *Notch1* show cyclic expression, including hairy and lunatic fringe (*Lfng*) [44,45,46]. *LFNG* expression in chicken embryos cycles every 90 minutes, defining a region just posterior to the last formed somite pair, specifying the following pair of somites. [44]. This cyclic signal is controlled through a negative feedback loop that controls the periodic expression of *Notch* and downstream effectors (**Fig. 8C)**. Studies using ubiquitous overexpression of *LFNG* lead to a truncated axis and failure of somitogenesis [44]. These data indicate that the cycling of LFNG is required for the proper periodic patterning of somites [44].

Only when cells of the presomitic mesoderm have moved anterior to the point where they have reached the determination front, and the oscillating clock of gene expression has reached them, do they differentiate to form epithelial somites (**Fig. 8D)** [1]. This can be seen with markers such as *MESO1*, which label the competent presomitic mesoderm, transitioning to form the next somite [47,48]. This expression also acts to down regulate expression of the presomitic marker, *TBX6* [1]. Only by continually elongating the axis does the source of the oscillation signals, the tailbud, move far enough away from the rostral presomitic mesoderm to continue to form new somites in a highly orcastrated series of signaling events.

Termination of axis elongation and somitogenesis

As the determination front moves posteriorly, cells are exposed to increasing levels of retinoic acid and decreasing levels of FGF. Premature exposure to retinoic acid, or down regulation of *Cyp26a1*, which metabolizes retinoic acid leads to a truncated axis through premature differentiation and down regulation of signals, such as WNT3A, which are required for tailbud progenitor maintenance [28,40,41,49,50]. The termination of somitogenesis occurs when the remainder of the tailbud progenitor population is exposed to retinoic acid and undergoes apoptosis [39,40]. The mechanism governing *Raldh2* expression within the tailbud has yet to be identified. Beginning at HH15, the presomitic mesoderm is reduced in size at each successive stage until the remaining tissue is approximately the same size as the last somite at HH25 [39]. It is likely that this continued shortening of the presomitic mesoderm results in exposure to retinoid signaling in the posterior of the embryo tailbud, terminating somitogenesis and removing any remaining cells through apoptosis [39].

Premature exposure to retinoic acid leads to a truncated axis through premature differentiation and down regulation of signals, such as WNT3A, which are required for tailbud precursor maintenance [28,40,41,49,50]. In addition, exposure to retinoic acid down regulates FGF signaling, which is followed by the down regulation of *TBX6,* brachyury, and *CYP26A1* **(Fig. 9A-N)** [39,40]. Following down regulation of these signals, there is an increase in programmed cell death within the posterior most cells of the tailbud **(Fig.12 O-Q)** [14,39,40]. Ectopic exposure of the tail to retinoic acid caused

the tailbud population to undergo premature apoptosis, leading to a truncated axis [40,50]. The role of programmed cell death in the tailbud remains unclear, but is suggested to remove any remaining undifferentiated tailbud cells at the end of somitogenesis **(Fig. 9R)** [14,39,40].

Studies suggest that although cells contributing to secondary body axis formation are proliferating, the most caudal tailbud population becomes non-proliferative [10,14]. It is unclear what level of proliferation is required for continued elongation of the tail. Furthermore, little work has been done to quantify cellular proliferation in the cell populations of the tail during the end of somitogenesis.

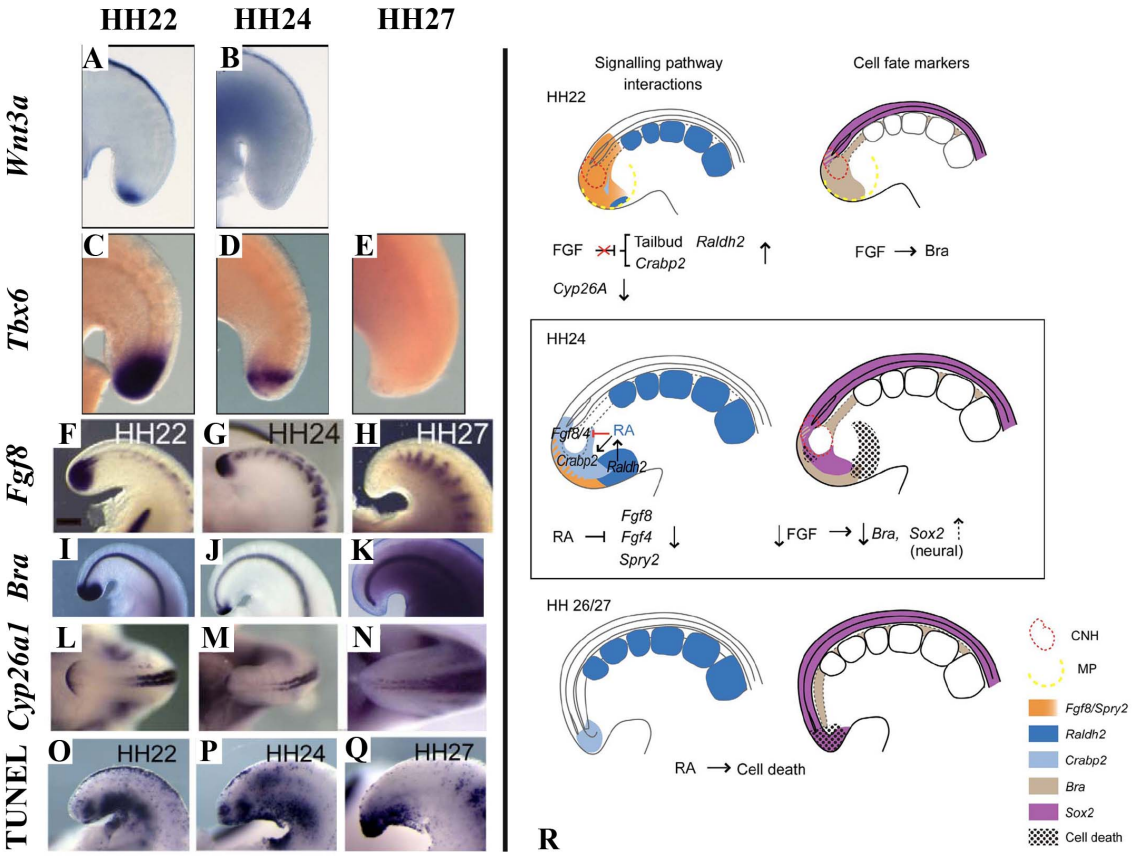

HH22
HH24
HH27
Wnt3a
Tbx6
Fgf8
Bra
Cyp26a1
TUNEL
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
HH22
HH24
HH27
Signalling pathway interactions
Cell fate markers
HH22
FGF
Tailbud
Raldh2
Crabp2
Cyp26A
FGF → Bra
HH24
Fgf8/4
RA
Crabp2
Raldh2
RA
Fgf8
Fgf4
Spry2
FGF
Bra,
Sox2
(neural)
HH 26/27
RA → Cell death
CNH
MP
Fgf8/Spry2
Raldh2
Crabp2
Bra
Sox2
Cell death
R

Figure 9: Termination of axis elongation and somitogenesis.

A-N: Somitogenesis in chicken embryos ends between HH24-27, and is marked by the down regulation of several genes involved in maintenance of the tailbud progenitor population or specifying mesoderm. These include *WNT3A* (A-B), *TBX6* (C-E), *FGF8* (F-H), *BRA* (I-K), and *CYP26A1* (L-N).

O-Q: Apoptosis is widespread during axis elongation, but becomes specific to the posterior most tailbud at the end of somitogenesis. Labeled by TUNEL.

A-E adapted from Tenin et al., 2010 [39]. Dorsal right, anterior up.

F-Q adapted from Olivera-Martinez et al., 2012 [40]. Anterior to right. A-K and O-Q sagittal view, L-N dorsal view.

R: The current model of the signals involved in the end of somitogenesis. Beginning at HH22, FGF no longer inhibits retinoic acid production or signaling, followed by a decrease in *CYP26A1*. By HH24 retinoic acid is expressed in the tailbud, which begins to inhibit FGF signaling. At HH27 FGF signaling is completely down regulated, and the remaining posterior most cells are undergoing apoptosis, with no more somites formed. Image from Olivera-Martinez et al., 2012 [40]. MP-mesoderm precursors.

Pathologies involving axis elongation and patterning

The term caudal dysplasia refers to varying degrees of developmental disorders involving the lumbar spinal column, sacrum and/or pelvis. These disorders are congenital, with the structures being malformed before birth. An example of a failure to form the correct number of vertebrae in humans is caudal regression syndrome. Individuals with caudal regression syndrome fail to form the caudal most vertebrae, including the coccyx (**Fig. 10A)** [51,52,53]. In addition, those with caudal regression syndrome often have lower limb deformities and deficiencies in innervation of the bladder and lower digestive tract [51,52]. The etiology and pathogenic mechanisms of such abnormalities have not been fully elucidated [2].

Genetic factors and teratogens are considered to be important predisposing elements. This includes teratogens such as retinoid and predisposing factors such as maternal diabetes [54,55]. Caudal regression syndrome has an estimated incidence of 1:25,000 to 1:60,000 births, however, in women with a maternal history of insulin-dependent diabetes mellitus, there is a 200 fold increased chance of having a child with caudal regression syndrome [51,52,56]. Mutations in the genes *Hlxb9, WNT3a*, *T-Brachyury* and *Lfng* have been identified as causative of axis truncation in humans and animal models [2,32,37,57,58,59,60]. However, despite this body of knowledge, many caudal dysplasias have no identified etiology or pathogenic mechanism.

Much of the knowledge regarding the mechanisms surrounding axis truncation has come from animal models of axis truncation, such as the T-brachyury mouse, the vestigial tail mouse, and zebrafish with a mutation in the no tail (*NTL*) gene (**Fig. 10B-C**) [32,35,59]. It is logical then, that the study of additional models of axis truncation could yield novel genes and pathways required for proper axis elongation. The Araucana breed of chicken is a model of axis truncation, as it lacks the caudal vertebrae and associated soft tissue **(Fig. 10D)**. As the dominant rumpless phenotype in chickens has not been studied since 1942, I propose to use Araucana as a model to study the genetics and morphogenesis responsible for axis elongation.

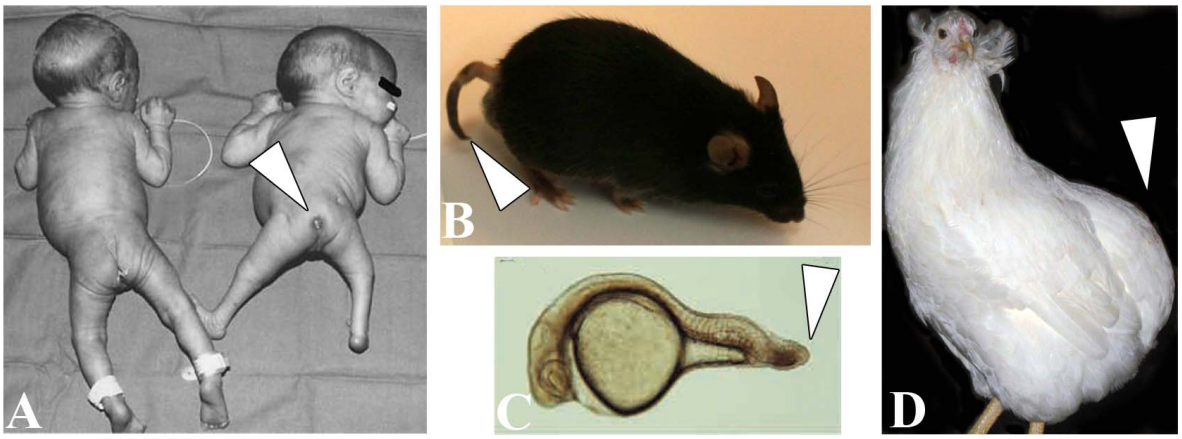


Figure 10: Examples of axis truncation.

A: Newborn monozygotic twins. The child on the right has caudal regression syndrome (white arrowhead). Image adapted from Zaw and Stone, 2002 [51].

B: Mouse heterozygous for *T* mutation (T11J). Note shortened, kinked tail (white arrowhead). Image from http://mousemutant.jax.org/images/nm4509T11Jpost.jpg. Accessed July 26, 2013. Image courtesy of The Jackson Laboratory.

C: Zebrafish embryo with mutation in *NTL*. Note the shortened and kinked tail (white arrowhead). Image from Amacher et al., 2002 [59]

D: Araucana rumpless chicken with the breed specific rumpless phenotype. Note the complete absence of the tail and tail feathers (white arrowhead). Image courtesy of Fritz Ludwig, Araucana Club of America.

Origin and breed characteristics of the Araucana chicken

The Araucana is a breed of the species *Gallus gallus domesticus*. The origin of Araucana has raised some debate, but evidence suggests an origination in Chile. Carbon dating analysis of chicken bone samples from Chile suggests that chickens may have been first introduced to South America from Polynesia. In addition, when the same chicken bones were genotyped, their haplotypes matched haplotypes of the current Chilean Araucana, suggesting Araucana originated from Asia and or Polynesia [61]. However, comparison of mitochondrial DNA haplotypes of these same samples revealed that they cluster more closely with chickens of a European descent [62], leading to the controversy surrounding the origin of the birds.

The North American breed standard of Araucana arose from two separate breeds, the Colloncas, which is rumpless and lays blue eggs, and the Quetero, which is tufted. Araucana were first imported into North America circa 1925. However, a breed standard was not determined until 1976, when the American Poultry Association first officially recognized Araucana as a breed. The American Poultry Association defines Araucana as both rumpless and tufted. In addition, Araucana are also known for laying blue eggs. As of March 2013, the Online Mendelian Inheritance in Animals database listed 206 known phenotypes in chicken, of which Araucana have three, rumplessness, tuftedness, and blue eggs (http://omia.angis.org.au/) **(Fig. 11A-D)**. Of these, only the blue egg phenotype has been identified to the level of the causative mutation. No mutation or mechanism has previously been described for the tufted or rumpless phenotypes.

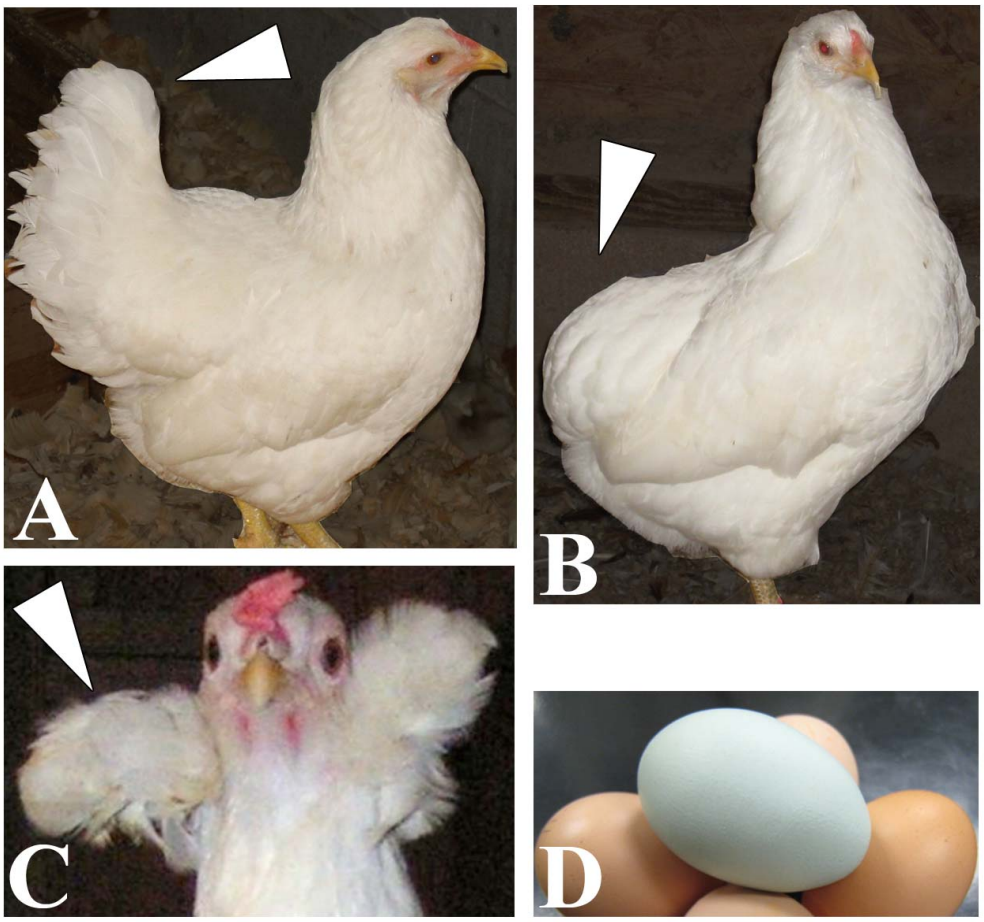


Figure 11: Comparison of the Araucana breed characteristics.

A: Araucana with a full tail (white arrowhead) and clean (no tufts) face.

B: Araucana with no tail (rumpless-white arrowhead) and clean face.

C: Araucana with bilateral tufts (white arrowhead) Image courtesy of Fritz Ludwig, Araucana Club of America.

D: Blue shelled Araucana egg compared to brown eggs.

Tufts are feather-covered peduncles that protrude from the side of the head (both sides or unilaterally) [63] **(Fig. 11C)**. Tuftedness (*Et*), or ear-tufts, is autosomal dominant mutation with reduced penetrance of 4-14% [63,64]. Based on inheritance studies, *Et* is homozygous lethal, with homozygous embryos dying before hatching [63,64]. The peduncle in Araucana^{Et} near the ear canal is thought to arise from failure of the fusion of the hyomandibular arches, however, no studies have been performed to further characterize the *Et* phenotype, or to identify candidate regions or genes associated with the *Et* locus [65].

Rumpless chickens lack the caudal vertebrae and associated soft tissue **(Fig 11. A-B)**. This includes the absence of the free caudal vertebrae, the fused vertebrae of the pygostyle, and in some cases 1-2 missing synsacral vertebrae [66]**.** Rumplessness is the result of an unidentified autosomal dominant mutation (Rp) [67,68,69]. Homozygosity for the rumpless trait is not embryonic lethal, and does not lead to increased embryonic mortality, however, chickens that are homozygous rumpless exhibit reduced fecundity [68].

As no studies on dominant rumplessness in chickens have been carried out since 1942 and no modern molecular approaches have been used to characterize the events of rumpless embryogenesis. The rumpless phenotype arises during early embryogenesis. The lack of the caudal vertebrae observed in preparations of rumpless skeletons suggests a defect in the embryonic formation of the caudal somites [66]. The two competing

hypotheses are that the somites were degraded after forming, or the full number of somites never formed [69]. More evidence exists for the latter hypothesis; however, no further studies have been done to determine the mechanism [69]. Furthermore, it is unclear what role the chordoneural hinge, ventral ectodermal ridge, cell fate identity, proliferation and apoptosis may play in the malformation observed in the Araucana^{Rp} phenotype. By identifying morphological processes involved in the formation of the Araucana^{Rp} phenotype during embryogenesis, we will better understand their roles in secondary body formation and somitogenesis.

Identifying candidate region(s) associated with a phenotype

As the mutation responsible for the rumpless phenotype in chickens is unknown, I propose to identify candidate mutations associated to the rumpless phenotype. By identifying genetic factors responsible for the Araucana^{Rp} phenotype, we will have a better understanding of the mechanism responsible for controlling axis length in vertebrates, and will potentially identify new gene targets in pathologies of axis elongation. Identification of the causative mutation(s)/inheritable factor(s) will also provide a starting point to better understand how the rumpless phenotype arises, and potentially identify novel pathways required for proper axis elongation.

A method to identify candidate genomic regions associated with a phenotype is a genome wide association study using single nucleotide polymorphisms (SNPs). SNPs are single base pair differences that occur between corresponding genetic loci or genes of an

individual, or between the DNA of two individuals of a species, resulting in genetic variation [70] **(Fig. 12A)**. Alternate versions of a genetic locus are called alleles. The combination of alleles, that are co-inherited, constitute a haplotype (**Fig. 12A)**. Haplotypes can change through additional mutations, or through genetic recombination during gametogenesis. During recombination, SNPs closest to an allele will more often remain with that allele than SNPs further away, which are more likely to recombine. Thus, the likelihood of recombination events between any two SNPs on a strand of DNA decreases the closer they are to each other. This leads to association of SNPs with alleles that are physically closer together, and therefore inherited together, and is referred to as a linkage disequilibrium (LD) [70]**.** When comparing the SNP profiles of an affected and unaffected population, individuals with a specific allele (causing a known phenotype) will have an increased chance of sharing a haplotype on the same chromosome, than unaffected individuals. Therefore, using a genome wide association study of SNPs to identify shared haplotypes amongst individuals with a shared phenotype allows for the identification of SNPs associated with the disease, and indicated the candidate genomic region(s) associated with that phenotype **(Fig. 12B)** [71,72,73,74]. This approach can only identify the region of interest and may contain multiple genes. The causative mutation, be it in the coding or regulatory region of a gene then has to be identified by further analysis.

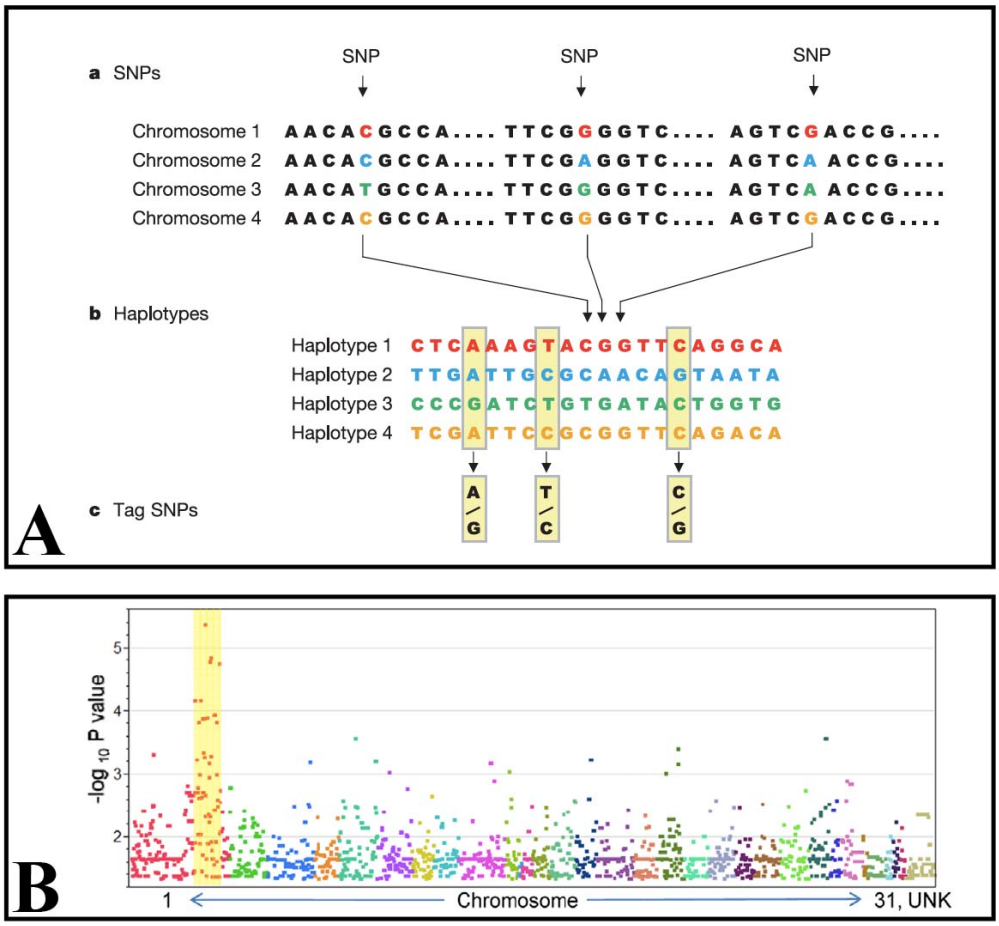


Figure 12: SNPs and haplotyping.

A: An illustration of various SNPs and haplotypes. (a) SNPs identified on the same chromosome between 4 individuals, with each SNP having two possible alleles. (b) Multiple SNPs with varying alleles across each chromosome is compared, with each chromosome having one of four haplotypes. (c) By examining a few of the SNPs, all four haplotypes can be determined. Image from The International HapMap Project, 2003 [70].

B: Manhattan plot, illustrating the use of SNP genotyping to identify candidate regions associated with a phenotype. Region of significance highlighted in yellow. Adapted from Brooks et al., 2010 [71].

The completion of the first draft chicken genome in 2004, and the subsequent analysis of breed variation yielded an initial map of 2.8 million SNPs [75,76]. Current maps have been expanded to include over 7 million reported SNPs [77], approximately 5 SNPs per kilobase, making SNPs a useful genomic marker [75] [77]. A SNP chip has been designed for high throughput sequencing of SNPs across the chicken genome [78,79].

Once a region associated with a phenotype is identified, candidate genes within that region can be studied. Sequencing is then performed to identify potential coding sequence mutations. However, inheritable factors causative for a phenotype are not required to be within coding sequences. Mutations are also found within regulatory regions such as promoters and enhancers, which are distributed across large genomic regions within vertebrate genomes. The relatively large intergenic regions make it challenging to identify individual mutations. Densities for SNP chips vary, with the 60k SNP chip having on average 5-30 kb between SNPs, with haplotype blocks often being much larger [73,79]. In order to sequence these larger blocks, either enrichment capture array and next generation sequencing is required, or whole genome resequencing [80].

No studies have been performed to identify the genetic mutation responsible for dominant rumplessness in chickens, nor has any current molecular methodology been applied to characterizing rumpless embryogenesis. *The overall goal of my research is to understand the genetic and molecular mechanisms required for axis elongation by identifying candidate mutations responsible for the Araucana^{Rp} phenotype, as well as characterizing morphogenesis during Araucana^{Rp} axis elongation.* I proposed three objectives centered on elucidating the genetic cause and developmental course through which the rumpless phenotype in Araucana^{Rp} arises.

1) **Identification of a candidate region(s) associated with the rumpless phenotype**

2) **Morphogenesis processes in axis elongation**

3) **Signaling and cell cycling processes in axis elongation**

References

1. Dequeant M-L, Pourquie O (2008) Segmental patterning of the vertebrate embryonic axis. Nat Rev Genet 9: 370-382.

2. Maroto M, Bone RA, Dale JK (2012) Somitogenesis. Development 139: 2453-2456.

3. Brent AE, Tabin CJ (2002) Developmental regulation of somite derivatives: muscle, cartilage and tendon. Current Opinion in Genetics & Development 12: 548-557.

4. Couly GF, Coltey PM, Le Douarin NM (1993) The triple origin of skull in higher vertebrates: a study in quail-chick chimeras. Development 117: 409-429.

5. Gaunt SJ, Dean W, Sang H, Burton RD (1999) Evidence that Hoxa expression domains are evolutionarily transposed in spinal ganglia, and are established by forward spreading in paraxial mesoderm. Mech Dev 82: 109-118.

6. Christ B, Huang R, Scaal M (2004) Formation and differentiation of the avian sclerotome. Anat Embryol (Berl) 208: 333-350.

7. Gomez C, Ozbudak EM, Wunderlich J, Baumann D, Lewis J, et al. (2008) Control of segment number in vertebrate embryos. Nature 454: 335-339.

8. Richardson MK, Allen SP, Wright GM, Raynaud A, Hanken J (1998) Somite number and vertebrate evolution. Development 125: 151-160.

9. Hamburger V, Hamilton LH (1951) A series of normal stages in the development of the chick embryo. Journal of Morphology 88: 49-92.

10. Schoenwolf GC (1977) Tail (End) bud contributions to the posterior region of the chick embryo. The Journal of Experimental Zoology 201: 227-245.

11. Schoenwolf GC (1979) Histological and ultrastructural observations of tail bud formation in the chick embryo. The Anatomical record 193: 131-148.

12. Ohta S, Suzuki K, Tachibana K, Tanaka H, Yamada G (2007) Cessation of gastrulation is mediated by suppression of epithelial-mesenchymal transition at the ventral ectodermal ridge. Development 134: 4315-4324.

13. McGrew MJ, Sherman A, Lillico SG, Ellard FM, Radcliffe PA, et al. (2008) Localised axial progenitor cell populations in the avian tail bud are not committed to a posterior Hox identity. Development 135: 2289-2299.

14. Mills CL, Bellairs R (1989) Mitosis and cell death in the tail of the chick embryo. Anatomy and Embryology 180: 301-308.

15. Catala M, Teillet M-A, Le Douarin NM (1995) Organization and development of the tail bud analyzed with the quail-chick chimaera system. Mechanisms of Development 51: 51-65.

16. Schoenwolf GC (1981) Morphogenetic processes involved in the remodeling of the tail region of the chick embryo. Anatomy and Embryology 162: 183-197.

17. Charrier JB, Teillet MA, Lapointe F, Le Douarin NM (1999) Defining subregions of Hensen's node essential for caudalward movement, midline development and cell survival. Development 126: 4771-4783.

18. Wilson V, Olivera-Martinez I, Storey KG (2009) Stem cells, signals and vertebrate body axis extension. Development 136: 1591-1604.

19. Cambray N, Wilson V (2002) Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. Development 129: 4855-4866.

20. Catala M, Teillet MA, De Robertis EM, Le Douarin ML (1996) A spinal cord fate map in the avian embryo: while regressing, Hensen's node lays down the notochord and floor plate thus joining the spinal cord lateral walls. Development 122: 2599-2610.

21. Colas JF, Schoenwolf GC (2001) Towards a cellular and molecular understanding of neurulation. Dev Dyn 221: 117-145.

22. Schoenwolf GC, Delongo J (1980) Ultrastructure of secondary neurulation in the chick embryo. American Journal of Anatomy 158: 43-63.

23. Naiche LA, Holder N, Lewandoski M (2011) FGF4 and FGF8 comprise the wavefront activity that controls somitogenesis. Proc Natl Acad Sci U S A 108: 4018-4023.

24. Boulet AM, Capecchi MR (2012) Signaling by FGF4 and FGF8 is required for axial elongation of the mouse embryo. Dev Biol 371: 235-245.

25. Martin BL, Kimelman D (2008) Regulation of canonical Wnt signaling by Brachyury is essential for posterior mesoderm formation. Developmental Cell 15: 121-133.

26. Yamaguchi TP, Takada S, Yoshikawa Y, Wu N, McMahon AP (1999) T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. Genes & Development 13: 3185-3190.

27. Takemoto T, Uchikawa M, Yoshida M, Bell DM, Lovell-Badge R, et al. (2011) Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. Nature 470: 394-398.

28. Takada S, Stark KL, Shea MJ, Vassileva G, McMahon JA, et al. (1994) Wnt-3a regulates somite and tailbud formation in the mouse embryo. Genes Dev 8: 174-189.

29. Nowotschin S, Ferrer-Vaquer A, Concepcion D, Papaioannou VE, Hadjantonakis AK (2012) Interaction of Wnt3a, Msgn1 and Tbx6 in neural versus paraxial mesoderm lineage commitment and paraxial mesoderm differentiation in the mouse embryo. Dev Biol 367: 1-14.

30. Chapman DL, Papaioannou VE (1998) Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. Nature 391: 695-697.

31. Clements D, Taylor HC, Herrmann BG, Stott D (1996) Distinct regulatory control of the *Brachyury* gene in axial and non-axial mesoderm suggests separation of mesoderm lineages early in mouse gastrulation. Mechanisms of Development 56: 139-149.

32. Chesley P (1935) Development of the short-tailed mutant in the house mouse. Journal of Experimental Zoology 70: 429-459.

33. Martin BL, Kimelman D (2012) Canonical Wnt Signaling Dynamically Controls Multiple Stem Cell Fate Decisions during Vertebrate Body Formation. Developmental Cell 22: 223-232.

34. Yoshikawa Y, Fujimori T, McMahon AP, Takada S (1997) Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm development in the mouse. Dev Biol 183: 234-242.

35. Greco TL, Takada S, Newhouse MM, McMahon JA, McMahon AP, et al. (1996) Analysis of the vestigial tail mutation demonstrates that Wnt-3a gene dosage regulates mouse axial development. Genes & Development 10: 313-324.

36. Aulehla A, Pourquié O (2010) Signaling Gradients during Paraxial Mesoderm Development. Cold Spring Harbor Perspectives in Biology 2.

37. Pourquie O (2011) Vertebrate segmentation: from cyclic gene networks to scoliosis. Cell 145: 650-663.

38. Delfini MC, Dubrulle J, Malapert P, Chal J, Pourquie O (2005) Control of the segmentation process by graded MAPK/ERK activation in the chick embryo. Proc Natl Acad Sci U S A 102: 11343-11348.

39. Tenin G, Wright D, Ferjentsik Z, Bone R, McGrew M, et al. (2010) The chick somitogenesis oscillator is arrested before all paraxial mesoderm is segmented into somites. BMC Developmental Biology 10: 24.

40. Olivera-Martinez I, Harada H, Halley PA, Storey KG (2012) Loss of FGF-Dependent Mesoderm Identity and Rise of Endogenous Retinoid Signalling Determine Cessation of Body Axis Elongation. PLoS Biol 10: e1001415.

41. Abu-Abed S, Dollé P, Metzger D, Beckett B, Chambon P, et al. (2001) The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. Genes & Development 15: 226-240.

42. Diez del Corral R, Olivera-Martinez I, Goriely A, Gale E, Maden M, et al. (2003) Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. Neuron 40: 65-79.

43. Dubrulle J, McGrew MJ, PourquiÈ O (2001) FGF Signaling Controls Somite Boundary Position and Regulates Segmentation Clock Control of Spatiotemporal Hox Gene Activation. Cell 106: 219-232.

44. Dale JK, Maroto M, Dequeant ML, Malapert P, McGrew M, et al. (2003) Periodic notch inhibition by lunatic fringe underlies the chick segmentation clock. Nature 421: 275-278.

45. Palmeirim I, Henrique D, Ish-Horowicz D, Pourquie O (1997) Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. Cell 91: 639-648.

46. Aulehla A, Wiegraebe W, Baubet V, Wahl MB, Deng C, et al. (2008) A beta-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. Nat Cell Biol 10: 186-193.

47. Yoon JK, Moon RT, Wold B (2000) The bHLH class protein pMesogenin1 can specify paraxial mesoderm phenotypes. Dev Biol 222: 376-391.

48. Yoon JK, Wold B (2000) The bHLH regulator pMesogenin1 is required for maturation and segmentation of paraxial mesoderm. Genes Dev 14: 3204-3214.

49. Iulianella A, Beckett B, Petkovich M, Lohnes D (1999) A Molecular Basis for Retinoic Acid-Induced Axial Truncation. Developmental Biology 205: 33-48.

50. Shum ASW, Poon LLM, Tang WWT, Koide T, Chan BWH, et al. (1999) Retinoic acid induces down-regulation of Wnt-3a, apoptosis and diversion of tail bud cells to a neural fate in the mouse embryo. Mechanisms of Development 84: 17-30.

51. Zaw W, Stone DG (2002) Caudal regression syndrome in twin pregnancy with type II diabetes. Journal of perinatology 22: 171-174.

52. Singh SK, Singh RD, Sharma A (2005) Caudal regression syndrome-case report and review of literature. Pediatric surgery international 21: 578-581.

53. Al Kaissi A, Klaushofer K, Grill F (2008) Caudal regression syndrome and popliteal webbing in connection with maternal diabetes mellitus: a case report and literature review. Cases Journal 1: 407.

54. Rothman KJ, Moore LL, Singer MR, Nguyen U-SDT, Mannino S, et al. (1995) Teratogenicity of High Vitamin A Intake. The New England Journal of Medicine 333: 1369-1373.

55. Chan BWH, Chan K-s, Koide T, Yeung S-m, Leung MBW, et al. (2002) Maternal Diabetes Increases the Risk of Caudal Regression Caused by Retinoic Acid. Diabetes 51: 2811-2816.

56. Subtil D, Cosson M, Houfflin V, Vaast P, Valat A, et al. (1998) Early detection of caudal regression syndrome: specific interest and findings in three cases. Eur J Obstet Gynecol Reprod Biol 80: 109-112.

57. Sparrow DB, Chapman G, Wouters MA, Whittock NV, Ellard S, et al. (2006) Mutation of the LUNATIC FRINGE gene in humans causes spondylocostal dysostosis with a severe vertebral phenotype. Am J Hum Genet 78: 28-37.

58. Ghebranious N, Blank RD, Raggio CL, Staubli J, McPherson E, et al. (2008) A Missense *T*(*Brachyury*) Mutation Contributes to Vertebral Malformations. Journal of Bone and Mineral Research 23: 1576-1583.

59. Amacher SL, Draper BW, Summers BR, Kimmel CB (2002) The zebrafish T-box genes no tail and spadetail are required for development of trunk and tail mesoderm and medial floor plate. Development 129: 3311-3323.

60. Belloni E, Martucciello G, Verderio D, Ponti E, Seri M, et al. (2000) Involvement of the HLXB9 homeobox gene in Currarino syndrome. Am J Hum Genet 66: 312-319.

61. Storey AA, Ramirez JM, Quiroz D, Burley DV, Addison DJ, et al. (2007) Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. Proc Natl Acad Sci U S A 104: 10335-10339.

62. Gongora J, Rawlence NJ, Mobegi VA, Jianlin H, Alcalde JA, et al. (2008) Indo-European and Asian origins for Chilean and Pacific chickens revealed by mtDNA. Proc Natl Acad Sci U S A 105: 10308-10313.

63. SOMES RG (1978) Ear-tufts: a skin structure mutation of the Araucana fowl. Journal of Heredity 69: 91-96.

64. Ralph G. Somes J, Pabilonia MS (1981) Ear tuftedness: a lethal condition in the Araucana fowl. The Journal of Heredity 72: 121-124.

65. Pabilonia MS, Somes RG, Jr. (1983) The embryonic development of ear-tufts and associated structural head and neck abnormalities of the Araucana fowl. Poult Sci 62: 1539-1542.

66. Landauer W, Dunn LC (1925) Two Types of Rumplessness in Domestic Fowls: A Morphological Comparison. Journal of Heredity 16: 153-160.

67. Dunn LC (1925) The inheritance of rumplessness in the domestic fowl. Journal of Heredity 16: 127-134.

68. Dunn LC, Landauer W (1934) The genetics of the rumpless fowl with evidence of a case of changing dominance. Journal of Genetics 29: 217-243.

69. Zwilling E (1942) The development of dominant rumplessness in chick embryos. Genetics 27: 641-656.

70. (2003) The International HapMap Project. Nature 426: 789-796.

71. Brooks SA, Gabreski N, Miller D, Brisbin A, Brown HE, et al. (2010) Whole-Genome SNP Association in the Horse: Identification of a Deletion in Myosin Va Responsible for Lavender Foal Syndrome. PLoS Genetics 6: e1000909.

72. Wragg D, Mwacharo JM, Alcalde JA, Hocking PM, Hanotte O (2012) Analysis of genome-wide structure, diversity and fine mapping of Mendelian traits in traditional and village chickens. Heredity (Edinb) 109: 6-18.

73. Robb EA, Gitter CL, Cheng HH, Delany ME (2011) Chromosomal mapping and candidate gene discovery of chicken developmental mutants and genome-wide variation analysis of MHC congenics. J Hered 102: 141-156.

74. Charlier C, Coppieters W, Rollin F, Desmecht D, Agerholm JS, et al. (2008) Highly effective SNP-based association mapping and management of recessive defects in livestock. Nature Genetics 40: 449-454.

75. Consortium ICPM (2004) A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. Nature 432: 717-722.

76. Consortium. ICGS (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432: 695-716.

77. Rubin C-J, Zody MC, Eriksson J, Meadows JRS, Sherwood E, et al. (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. Nature 464: 587-591.

78. Muir WM, Wong GK, Zhang Y, Wang J, Groenen MAM, et al. (2008) Review of the initial validation and characterization of a 3K chicken SNP array. World's Poultry Science Journal 64: 219-225.

79. Groenen M, Megens H-J, Zare Y, Warren W, Hillier L, et al. (2011) The development and characterization of a 60K SNP chip for chicken. BMC Genomics 12: 274.

80. Robb EA, Delany ME (2012) Case Study of Sequence Capture Enrichment Technology: Identification of Variation Underpinning Developmental Syndromes in an Amniote Model. Genes 3: 233-247.

CHAPTER 2

GENOME-WIDE ASSOCIATION MAPPING AND IDENTIFICATION OF CANDIDATE GENES FOR THE RUMPLESS AND EAR-TUFTED TRAITS OF THE ARAUCANA CHICKEN

Rooksana E. Noorai[1], Nowlan H. Freese[2], Lindsay M. Wright[1], Susan C. Chapman[2]*, Leigh Anne Clark[1]*

1 Department of Genetics and Biochemistry, Clemson University, Clemson, South Carolina, United States of America, **2** Department of Biological Sciences, Clemson University, Clemson, South Carolina, United States of America

*Co-corresponding authors

lclark4@clemson.edu

schapm2@clemson.edu

This manuscript was published online in *PLoS ONE* on July 23, 2012, and is in the required journal format.

Abstract

Araucana chickens are known for their rounded, tailless rumps and tufted ears. Inheritance studies have shown that the rumpless (*Rp*) and ear-tufted (*Et*) loci each act in an autosomal dominant fashion, segregate independently, and are associated with an increased rate of embryonic mortality. To find genomic regions associated with *Rp* and *Et*, we generated genome-wide SNP profiles for a diverse population of 60 Araucana chickens using the 60K chicken SNP BeadChip. Genome-wide association studies using 40 rumpless and 11 tailed birds showed a strong association with rumpless on Gga 2 (P_{raw} = 2.45 x 10^{-10}, P_{genome} = 0.00575), and analysis of genotypes revealed a 2.14 Mb haplotype shared by all rumpless birds. Within this haplotype, a 0.74 Mb critical interval containing two iroquois homeobox genes, *Irx1* and *Irx2*, was unique to rumpless Araucana chickens. *Irx1* and *Irx2* are central for developmental prepatterning, but neither gene is known to have a role in mechanisms leading to caudal development. A second genome-wide association analysis using 30 ear-tufted and 28 non-tufted birds revealed an association with tufted on Gga 15 (P_{raw} = 6.61 x 10^{-7}, P_{genome} = 0.0981). We identified a 0.58 Mb haplotype common to tufted birds and harboring 7 genes. Because homozygosity for *Et* is nearly 100% lethal, we employed a heterozygosity mapping approach to prioritize candidate gene selection. A 60 kb region heterozygous in all Araucana chickens contains the complete coding sequence for *TBX1* and partial sequence for *GNB1L*. *TBX1* is an important transcriptional regulator of embryonic development and a key genetic determinant of human DiGeorge syndrome. Herein, we describe localization of *Rp* and *Et* and identification of positional candidate genes.

Introduction

There are hundreds of domestic chicken breeds worldwide [1]. Breeds were generally developed for meat and egg production, but morphological traits, plumage color, and other distinctive characteristics were also selected. The Araucana chicken, originally from Chile, is a multi-purpose breed initially established for its blue-shelled eggs [1,2]. Araucana chickens are also known for two other distinguishing traits: a rounded, tailless rump and protruding ear-tufts. Although these traits segregate in the population, the United States Araucana breed standard requires show birds to possess both phenotypes.

The rumpless phenotype is characterized by the absence of all free caudal vertebrae and the uropygial gland [3]. Without underlying skeletal support, birds with caudal truncation lack a fleshy rump and tail feathers [3]. An intermediate rumpless phenotype, wherein some caudal vertebrae are present but irregularly fused together, is thought to result from a modifier gene introduced through crosses with non-Araucana tailed chickens [3,4]. The rumpless phenotype arises from a defect in caudal patterning that is controlled by a dominant gene (*Rp*) [3]. Rumpless Araucana chickens may be heterozygous or homozygous for this locus. In test matings, all rumpless intermediates were determined to be heterozygous (Rp/rp^{+}) [3]. Homozygosity is underrepresented among chicks from rumpless to rumpless matings, indicating that the *Rp*/*Rp* genotype has reduced viability [3,5]. Birds having at least one copy of *Rp* have increased mortality in the embryonic stage, with death occurring at 17 to 21 days of incubation [3]. Rumpless birds also have reduced fecundity as adults [3].

Ear-tufts are feather-covered, epidermal protrusions originating near the ear canal (Figure 1). The mass of tissue forming the protrusion, or peduncle, is believed to develop as a result of the incomplete fusion of the hyomandibular arches, and it can vary in position and length (from 2 mm to 2 cm) [6,7]. Tufted chickens may also have structural rearrangement of the ears [6]. Abnormalities include irregularly shaped external ear openings and shortened or absent external auditory canals [6].

Inheritance studies indicate that tufted is governed by a dominant locus, *Et* [6,8]. Test matings show that all tufted birds are heterozygous (Et/et^+) and that homozygosity for *Et* is lethal at about 17-19 days of incubation [6,8]. Lethality among a portion of heterozygous birds is also reported, appearing to occur at 20-21 days of incubation [8]. Post-hatch mortality is significantly higher among tufted chickens [6,8].

Because tufts can occur unilaterally or bilaterally and may differ in size from one side to the other, *Et* is proposed to have variable expressivity [6]. In addition, a paucity of tufted progeny from mating studies in 1978 suggests reduced penetrance of the tufted locus [6]. In 1981, Somes and Pabilonia identified a tufted male that produced excessive tufted progeny when crossed with an et^+/et^+ White Leghorn (86%), and they speculated that *Et*/*Et* birds may occasionally reach maturity [8]. The non-tufted chicks from the *Et*/*Et* male produced tufted progeny when crossed with an et^+/et^+ White Leghorn, indicating that their predicted genotype does not match their phenotype, providing further evidence for variable penetrance.

The aim of our investigation was to localize the genetic bases for the rumpless and tufted phenotypes of the Araucana chicken. To this end, we generated genome-wide SNP profiles for 60 Araucana chickens using the 60K chicken SNP BeadChip [9]. Using a genome-wide association approach, we elucidate the chromosomal regions harboring *Rp* and *Et* and identify strong candidate genes for each trait.

Results

Case/control analyses were carried out using 40 rumpless and 11 tailed Araucana chickens (Figure 2a). Seven birds described as having partial tails by their breeders were excluded from the rumpless association analysis because of uncertainty concerning their phenotype. A total of 191 SNPs were associated with the rumpless phenotype ($P_{raw} \leq 0.0001$), 72 of which were located on Gga 2 (Figure 2b). The most significant result obtained was for SNP Gga_rs13637596, located on chromosome 2 at position 88.95 Mb ($P_{raw} = 2.45 \times 10^{-10}$, $P_{genome} = 0.00575$). The next two most significant results were for proximal SNPs located at 89.17 Mb ($P_{raw} = 1.20 \times 10^{-9}$, $P_{genome} = 0.0119$) and 89.19 Mb ($P_{raw} = 1.20 \times 10^{-9}$, $P_{genome} = 0.0119$).

Analysis of genotypes in the Gga 2 region revealed a 2.14 Mb haplotype (87.99 – 90.13 Mb) predicted to contain five genes (Figure 3). All 40 rumpless birds had at least one copy of the haplotype: 18 were homozygous and 22 were heterozygous. Partial tailed birds were heterozygous. The haplotype was absent in its entirety from the 11 tailed birds. Three tailed birds were heterozygous for partial blocks of the haplotype and further delimit the critical interval to 0.74 Mb (88.77 - 89.51 Mb). This region contains two candidate genes: *Irx1* and *Irx2*.

Analyses for association with the tufted phenotype, using 30 cases and 28 controls, resulted in 31 significant SNPs, 11 of which map to Gga 15 (Figure 2c). The most significant results were for SNPs Gga_rs10730189 ($P_{raw} = 6.61 \times 10^{-7}$, $P_{genome} = 0.0981$)

and Gga_rs15762547 (P_{raw} = 9.19 x 10^{-7}, P_{genome} = 0.118), located at positions 1.33 Mb and 1.30 Mb on chromosome 15, respectively. Four other proximal SNPs also reached significance (Figure 2d).

Analysis of genotypes reveals that 29 of 30 tufted birds shared a haplotype extending from the telomere of Gga 15 to position 1.75 Mb. These birds were heterozygous for the complete haplotype. Two of 28 non-tufted birds were also heterozygous for the haplotype in its entirety. A single tufted bird shared only part of the 1.75 Mb haplotype, defining a 0.58 Mb (0.90 – 1.48 Mb) critical interval that is heterozygous in all 30 tufted birds and contains 7 genes. Because tufted is nearly always recessive lethal, blocks of homozygosity for the tufted haplotype were identified to reduce the number of candidate genes. Homozygosity blocks in three birds flank a 60 kb interval harboring two genes: *TBX1* and *GNB1L* (Figure 4).

Discussion

In this study, we used genome-wide SNP profiles to localize genes causative for two breed-defining phenotypes of Araucana chickens, rumpless and ear-tufts. We took advantage of the fact that both traits segregate independently in the population by using a single data set to carry out an association analysis for each trait. Haplotype analyses based on inheritance patterns were used to identify positional candidate genes for both traits.

We identified a rumpless haplotype spanning 2.14 Mb and five genes on chromosome 2. The haplotype is present in the heterozygous or homozygous state in rumpless birds. All 7 birds with partial tails are heterozygous for the rumpless haplotype and likely represent the intermediate phenotype described by Dunn and Landauer [3]. Because rumpless is dominant and fully penetrant, we further delimited the critical interval by identifying regions of the haplotype shared by tailed birds. A 0.74 Mb region common to all rumpless birds, and absent from 11 tailed birds, harbors *Rp*.

These data reveal that *Rp* maps to a region of Gga 2 that is distinct from the predicted location of genes previously associated with caudal truncation [10-14]. The 0.74 Mb critical interval contains the iroquois homeobox genes, *Irx1* and *Irx2.* The iroquois genes encode transcription factors that function in patterning and regionalization of tissues early in development [15]. *Irx1* and *Irx2* are prepattern and proneural genes first identified in *Drosophila* and *Xenopus* [16,17]. Studies of gene function suggest that *Irx* genes have

redundant yet distinct roles in development [18,19]. *Irx* genes have been knocked out in mice and zebrafish with little effect on tail development [19-23]. However, the rumpless phenotype is dominant, suggesting that misexpression of *Irx1* or *Irx2* may underlie the trait, rather than loss of function.

We identified SNPs on Gga 15 that are strongly associated with the tufted phenotype and define a 0.58 Mb haplotype for which all tufted birds in our cohort are heterozygous. No birds are homozygous for the complete tufted haplotype. These data support conclusions from previous inheritance studies that suggest nearly 100% of tufted birds are heterozygous, and that *Et*/*Et* is lethal [6,8].

Two non-tufted Araucana chickens are heterozygous for the tufted haplotype. These birds may signify reduced penetrance. Penetrance of the tufted allele is estimated to range from 86% to 96% [6,8]. Based on the assigned phenotypes and the associated haplotype, we observed 94% penetrance in our cohort. Alternatively, these birds may have been incorrectly phenotyped by their breeders due to short peduncles or missing protruding feathers.

The 0.58 Mb haplotype harbors 7 protein-coding genes. Unlike rumpless, identification of the tufted haplotype in non-tufted birds could not be used to narrow the critical interval because of reduced penetrance. However, because homozygosity for *Et* is nearly always lethal, we were able to prioritize candidate gene selection using heterozygosity mapping.

Tufted birds with blocks of homozygosity extending into the 0.58 Mb common haplotype were identified, and these regions were deemed less likely to harbor the *Et* locus. These data indicate that *Et* is located in a region containing partial coding sequence for *GNB1L*, which encodes a protein implicated in neuropsychiatric disorders [24,25], and complete coding sequence for *TBX1* [26], an important transcriptional regulator of embryonic development.

Haploinsufficiency for *TBX1* is considered to be the key genetic determinant of human DiGeorge syndrome (DGS), which is caused by a heterozygous chromosomal deletion of 22q11.2 [27]. While the clinical phenotype is highly variable, DGS is characterized by craniofacial and cardiovascular abnormalities. Malformations in DGS are attributed to disturbed segmentation and patterning of the pharyngeal structures [28]. Auricular defects common in DGS include narrow or absent external ear canal and protruding ears [29]. Homozygosity for null mutations of *TBX1* in mice and zebrafish causes a range of phenotypic effects similar to DGS, including abnormal ear development [30,31]. Based on phenotypic similarities between the malformations causing ear tufts and DGS, *TBX1* is a highly plausible candidate gene and the primary focus of ongoing work to identify the genetic basis for ear-tufts in Araucana chickens.

In conclusion, we used genome-wide association and haplotype analyses to localize *Rp* and *Et* to chicken chromosomes 2 and 15, respectively. In addition, we identified candidate genes that are immediate targets for future work.

Materials and Methods

Ethics Statement

This study was approved by the Clemson University IACUC protocol number 2011-041 and IBC protocol number 2010-041.

Study Cohort

Whole blood for DNA was collected from 6 different flocks of Araucana chickens from the United States. Phenotypic information and photographs, when available, were provided by owners. Birds with tufts of any size and on either side of the head were classified as tufted. Because both traits segregate in the Araucana population, birds were selected to ensure that the phenotypes were balanced. Our study cohort comprised 60 Araucana chickens: 21 rumpless/tufted birds, 20 rumpless/non-tufted birds, 7 tailed/non-tufted birds, 5 tailed/tufted birds, 5 partial/tufted birds, and 2 partial/non-tufted birds. Genomic DNA was isolated using the DNeasy blood and tissue kit (QIAGEN, Valencia, USA) and adjusted to a concentration of 50 ng/uL.

Genome-wide Association Mapping

SNP genotypes were generated using the Illumina 60K chicken SNP BeadChip, which has 57,636 SNPs across chromosomes 1 through 28, Z, W, and two unmapped linkage groups [9]. BeadChips were processed by DNA Landmarks (Quebec, Canada), according to manufacturer's protocols. Raw data files were analyzed using GenomeStudio's Genotyping Module to generate SNP calls. The PLINK Input Report Plug-in v2.1.1 was

used to format the data. For analysis, Gga 27, Gga 28, Gga Z, Gga W, and microchromosomes were all identified as chromosome zero. Case/control analyses using 56,685 SNPs were performed using PLINK [32]. Two birds with excessive missing data were excluded from all analyses. By convention, P_{raw} values ≤ 0.0001 were considered significant. Permutation testing, using 100,000 iterations, was carried out using PLINK.

Acknowledgements

We are grateful to the Araucana Club of America and their members who provided samples, the Morgan Poultry Center at Clemson University for their assistance, and the Clemson University Genomics Institute for use of software and hardware.

References

1. Ekarius C (2007) Storey's Illustrated Guide to Poultry Breeds. China: Storey Publishing. pp. 23-24.

2. Browman DL (1978) Advances in Andean Archaeology. Great Britain: Mouton Publishers. pp. 189-196.

3. Dunn LC, Landauer W (1934) The genetics of the rumpless fowl with evidence of a case of changing dominance. J Genet 29: 217-243.

4. Dunn LC, Landauer W (1936) Further data on genetic modification of rumplessness in the fowl. J Genet 33: 401-405.

5. Zwilling E (1942) The development of dominant rumplessness in chick embryos. Genetics 27: 641-656.

6. Somes Jr RG (1978) Ear-Tufts: a skin structure mutation of the Araucana fowl. J Hered 69: 91-96.

7. Pabilonia MS, Somes Jr RG (1983) The Embryonic Development of Ear-Tufts and Associated Structural Head and Neck Abnormalities of the Araucana Fowl. Poult Sci 62: 1539-1542.

8. Somes Jr RG, Pabilonia MS (1981) Ear tuftedness: a lethal condition in the Araucana fowl. J Hered 72: 121-124.

9. Groenen MAM, Megens HJ, Zare Y, Warren WC, Hillier LW, et al. (2011) The development and characterization of a 60K SNP chip for chicken. BMC Genomics 12: 274.

10. Herrmann BG, Labeit S, Poustka A, King TR, Lehrach H (1990) Cloning of the T gene required in mesoderm formation in the mouse. Nature 343: 617-622.

11. Greco TL, Takada S, Newhouse MM, McMahon JA, McMahon AP, et al. (1996) Analysis of the vestigial tail mutation demonstrates that Wnt-3a gene dosage regulates mouse axial development. Genes Dev 10: 313-324.

12. Ross AJ, Ruiz-Perez V, Wang Y, Hagan DM, Scherer S, et al. (1998) A Homeobox gene, HLXB9, is the major locus for dominantly inherited sacral agenesis. Nat Genet 20: 358-361.

13. Abu-Abed S, Dollé P, Metzger D, Beckett B, Chambon P, et al. (2001) The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. Genes Dev 15: 226-240.

14. van den Akker E, Forlani S, Chawengsaksophak K, de Graaff W, Beck F, et al. (2002) *Cdx1* and *Cdx2* have overlapping functions in anteroposterior patterning and posterior axis elongation. Development 129: 2181-2193.

15. Cavodeassi F, Modolell J, Gómez-Skarmeta JL (2001) The Iroquois family of genes: from body building to neural patterning. Development 128: 2847-2855.

16. Gómez-Skarmeta JL, Diez del Corral R, de la Calle-Mustienes E, Ferré-Marcó D, Modolell J (1996) Araucan and caupolican, two members of the novel iroquois complex, encode homeoproteins that control proneural and vein-forming genes. Cell 85: 95-105.

17. Gómez-Skarmeta JL, Modolell J (1996) araucan and caupolican provide a link between compartment subdivisions and patterning of sensory organs and veins in the Drosophila wing. Genes Dev 10: 2935-2945.

18. Costantini DL, Arruda EP, Agarwal P, Kim KH, Zhu Y, et al. (2005) The homeodomain transcription factor Irx5 establishes the mouse cardiac ventricular repolarization gradient. Cell 123: 347-358.

19. Lebel M, Agarwal P, Cheng CW, Kabir MG, Chan, TY (2003) The Iroquois homeobox gene Irx2 is not essential for normal development of the heart and midbrain-hindbrain boundary in mice. Mol Cell Biol 23: 8216-8225.

20. Itoh M, Kudoh T, Dedekian M, Kim CH, Chitnis AB (2002) A role for *iro1* and *iro7* in the establishment of an anteroposterior compartment of the ectoderm adjacent to the midbrain-hindbrain boundary. Development 129: 2317-2327.

21. Peters T, Ausmeier K, Dildrop R, Rüther U (2002) The mouse Fused toes (Ft) mutation is the result of a 1.6-Mb deletion including the entire Iroquois B gene cluster. Mamm Genome 13: 186-188.

22. Cheng CW, Yan CHM, Hui CC, Strähle U, Cheng SH (2006) The Homeobox gene *irx1a* is required for the propagation of the neurogenic waves in the zebrafish retina. Mech Develop 123: 252-263.

23. Kimura W, Machii M, Xue X, Sultana N, Hikosaka K, et al. (2011) Irx1 mutant mice show reduced tendon differentiation and no patterning defects in musculoskeletal system development. Genesis 49: 2-9.

24. Williams NM, Glaser B, Norton N, Williams H, Pierce T, et al. (2008) Strong evidence that *GNB1L* is associated with schizophrenia. Hum Mol Genet 17: 555-566.

25. Li Y, Zhao Q, Wang T, Liu J, Li J, et al. (2011) Association study between *GNB1L* and three major mental disorders in Chinese Han populations. Psychiat Res 187: 457-459.

26. Völker M, Backström N, Skinner BM, Langley EJ, Bunzey SK, et al. (2010) Copy number variation, chromosome rearrangement, and their association with recombination during avian evolution. Gen Res 20: 503-511.

27. Yagi H, Furutani Y, Hamada H, Sasaki T, Asakawa S, et al. (2003) Role of *TBX1* in human del22q11.2 syndrome. Lancet 362: 1366-1373.

28. Wurdak H, Ittner LM, Sommer L (2006) DiGeorge syndrome and pharyngeal apparatus development. BioEssays 28: 1078-1086.

29. Butts SC (2009) The facial phenotype of the velo-cardio-facial syndrome. Int J Pediatr Otorhinolaryngol 73: 343-350.

30. Jerome LA, Papaioannou VE (2001) DiGeorge syndrome phenotype in mice mutant for the T-box gene, *Tbx1*. Nat Genet 27: 286-291.

31. Piotrowski T, Ahn DG, Schilling TF, Nair S, Ruvinsky I, et al. (2003) The zebrafish *van gogh* mutation disrupts *tbx1*, which is involved in the DiGeorge deletion syndrome in humans. Development 130: 5043-5052.

32. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559-575.

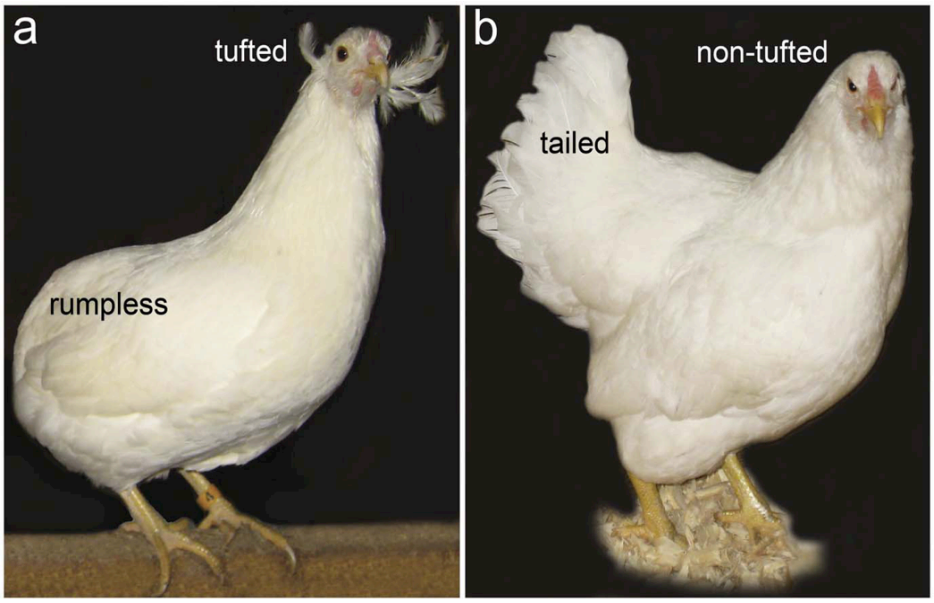


Figure 1. Araucana chicken. (a) General appearance of a rumpless, tufted Araucana chicken. **(b)** For comparison, a tailed, non-tufted Araucana chicken.

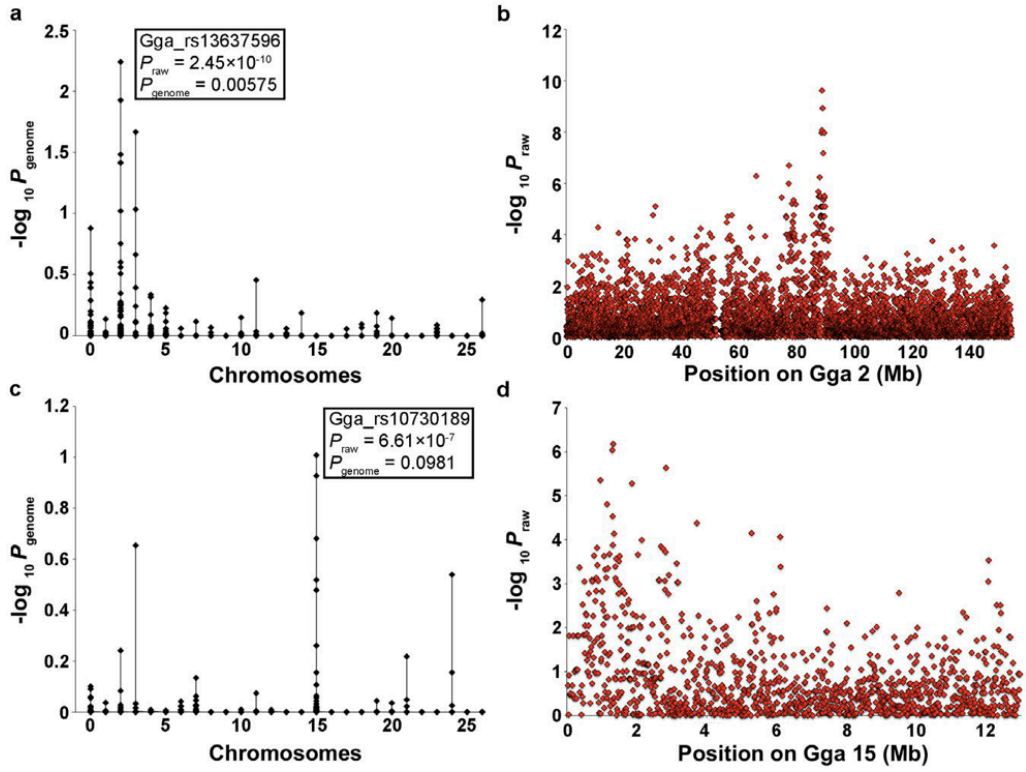


Figure 2. Genome-wide association for *Rp* and *Et*. After 100,000 permutations, the genome-wide adjusted *P* values ($-\log_{10} P_{genome}$) for each SNP are plotted by chromosome (left). The raw *P* values for the most strongly associated chromosomes are plotted against chromosomal position (right). **(a,b)** 40 rumpless versus 11 tailed Araucana chickens **(c,d)** 30 tufted versus 28 non-tufted Araucana chickens.

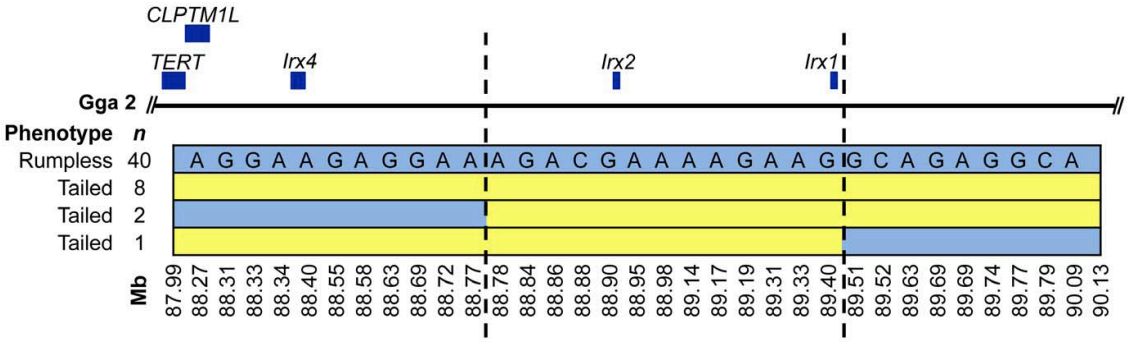


Figure 3. Localization of *Rp*. Physical map showing the relative positions of mapped genes and informative SNP markers within the 2.14 Mb rumpless haplotype on Gga 2. Light blue shading denotes the rumpless haplotype (alleles are shown in the top row). Dashed lines flank the critical interval wherein no tailed birds share the rumpless haplotype.

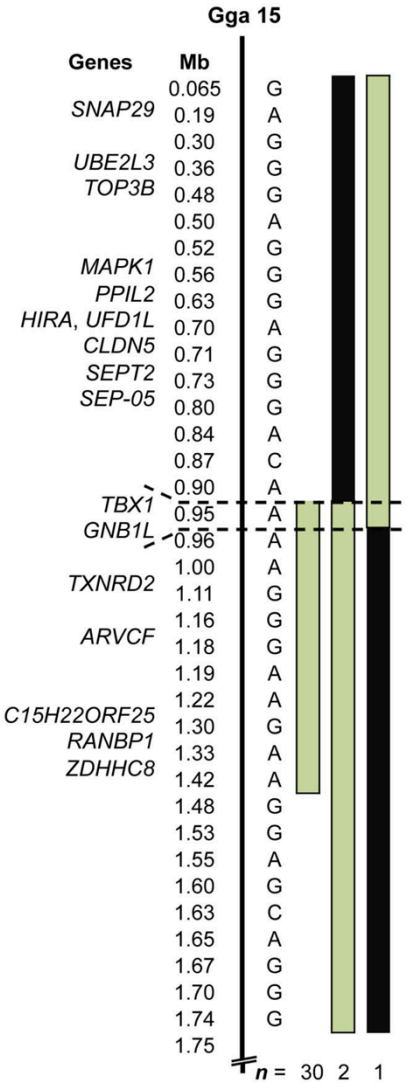


Figure 4. Localization of *Et*. Physical map showing the relative positions of genes and informative SNP markers in the associated region of Gga 15. Alleles of the tufted haplotype and positions are shown. Pale green bars denote heterozygosity for the tufted haplotype. Black bars denote homozygosity for the tufted haplotype. Dashed lines mark a 60 kb interval wherein all tufted birds are heterozygous for the haplotype.

CHAPTER THREE

A NOVEL GAIN OF FUNCTION OF THE *IRX1* AND *IRX2* GENES DISRUPTS AXIS ELONGATION IN THE ARAUCANA RUMPLESS CHICKEN

Nowlan H. Freese, Miguel Maroto, Brianna A. Lam, Allison Scott, and Susan C. Chapman

Clemson University, Department of Biological Sciences,132 Long Hall, Clemson, SC 29634-0314, USA

*Author for correspondence: schapm2@clemson.edu

Tel: 864-656-5432

Fax: 864-656-0435

Key words: Araucana, chick, development, Iroquois, rumpless, tail organizer, tailbud, somitogenesis, axis elongation.

Grant Sponsor: NIH Grant number: R01DC009236

This manuscript will be submitted to *PLoS GENETICS* and is in the required journal format.

Abstract

The molecular mechanism of the rumpless phenotype in the Araucana chicken has not been determined. We show that Araucana fail to form the free caudal vertebrae and pygostyle due to a lack of formation of the caudal embryonic somites. Sequencing of the 0.74Mb critical region to which the rumpless phenotype was previously mapped was carried out on six Araucana (3 homozygous rumpless, 2 heterozygous rumpless, and 1 homozygous tailed). We identified 298 candidate mutations, centered within a predicted regulatory region for proneural genes, Iroquois 1 (*Irx1*) and Iroquois 2 (*Irx2*). Using in situ hybridization (ISH) we show that both genes are misexpressed in the developing Araucana embryo tailbud. To identify pathways altered following misexpression of *Irx1* and *Irx2*, we have carried out a screen on genes required for proper caudal development using ISH. *Brachyury* and *Tbx6,* required for specification of mesoderm, are both down regulated. This is concurrent with down regulation of *Cyp26a1*, which protects the caudal tailbud from the anterior gradient of retinoic acid. Next, there is down regulation of *Fgf8* and *Wnt3a*, which are required for the maintenance of the tailbud progenitor population, as well as an increase in ectopic neural tissue within the tailbud. Furthermore, there is an increase in the level of apoptosis in the tailbud that could be removing the remaining tailbud progenitor population, and a concurrent decrease in proliferation. These results suggest that the shortened axis in rumpless Araucana chickens is due to a mutation that induces the aberrant expression of both *Irx1* and *Irx2* in the tailbud territory. This misexpression alters the fate of the tailbud progenitor population, which adopts a neural fate and/or is removed by apoptosis.

Introduction

Axis elongation of the vertebrate embryo is a complex process that involves the generation of different cell lineages from a stem cell progenitor population located in the tailbud region. Among their derivatives are the mesodermal cells that make the paraxial mesoderm (PM), a tissue from which epithelial somites are formed [1]. Somites are transitory structures that mature to generate the axial skeleton, including the vertebrae, cartilage, and most of the skeletal musculature and dermis. Somite formation or somitogenesis is a highly regulated process that operates under the control of an oscillatory mechanism known as the segmentation clock. The number of somites formed is species-specific and highly variable; for example whereas chickens have between 51-53 somites, corn snakes have approximately 315 somites [2] [3]. Current models of axis length termination include cessation of cell ingression into the tailbud through the ventral ectodermal ridge (VER), elimination of the tailbud progenitor population through apoptosis, and diminution of the presomitic mesoderm (PSM) leading to exposure to the differentiating effects of retinoic acid (RA) [4] [5] [6].

During vertebrate development, the axial elongation process can be divided into two parts, defined by the origin and anatomical location of the progenitor cells. During primary body formation, progenitor cells originate by gastrulation or continuous ingression of new cells in a territory known as the primitive streak and Hensen's node [7]. By HH14-15, during the late gastrula stage, the primitive streak and Hensen's node are replaced by a bulblike structure, the tailbud, consisting of a morphologically uniform

mass of mesenchyme [7] [8] [9]. Elongation of the tail and caudal somitogenesis occur during secondary body formation. The node streak border (NSB) and a portion of the caudal lateral epiblast (CLE) give rise to the chordo-neural hinge (CNH) [10,11]. The CNH population is located adjacent to the caudal end of the neural tube and notochord where they become indistinguishable from the surrounding mesoderm. There are currently no markers specific to the CNH. CNH progenitor cells give rise to the ventral cells of the medullary cord, which undergoes cavitation to form the neural tube, somite progenitors in the PSM, and caudal extension of the notochord [10] [12] [9]. The cells of the CNH are referred to as long-term axial progenitors, and together with the dorso-posterior tail bud and ventral tail bud populations (altogether the tailbud progenitor population) give rise to all the derivatives of the tail [12].

In contrast to the CNH, the dorso-posterior population produces only mesodermal derivatives, traced to the medial somitic tissues, and ventral tail bud cells, the lateral somite compartment [13] [14] [15] [16] [11]. In addition, following regression, the primitive streak and node give rise to the VER, which is the thickened ectodermal tissue located at the tailbud ventro-distally. Posterior epiblast cells undergo an epithelial to mesenchymal transition, ingressing through the VER to contribute to the tail bud mesenchyme (TBM) [17] [4].

Somitogenesis and axis elongation are critical processes for the proper morphogenesis of the embryo, and any interference brought about by teratogenic/environmental factors or

the existence of congenital mutations can lead to the generation of deformities. Many model organisms have been used to study axis elongation and associated defects including chicken, mouse, and zebrafish [18]. Notably in mice, the spontaneous mutants, the *vestigial tail* mouse, hypomorphic for *Wnt3a*, and the *Brachyury* mouse have been well studied as models of axis truncation [19] [20] [21] [22]. In these models, loss or decrease in expression of either *Wnt3a* or *T-brachyury* within the tailbud leads to a failure of maintenance of the tailbud progenitor population and of mesoderm specification. In addition to these models, many other mouse models of axis truncation have been studied through targeted gene knockout including *Fgf8* and *Cyp26a1* [18] [23] [24].

Araucana chickens have been maintained for their rumplessness and ear tufts [25,26] [27] [28] [29]. The rumpless gene (Rp) is inherited autosomal dominantly [27]. However, despite the popularity of rumpless Araucana (Araucana^{Rp}), the mechanism causing rumplessness remains unknown. Recently, we have mapped the rumpless mutation to a 0.74 Mb region on chromosome 2 [30]. However, candidate mutations have not been identified. Thus, the rumpless mutation in the Araucana breed offers an opportunity to study the morphological and molecular mechanisms of tail development as well as the cessation of axis elongation in an accessible model organism. Our aim is to identify and molecularly characterize pathways responsible for axis truncation in rumpless Araucana.

In the current study we show that caudal truncation in Araucana^{Rp} is due to a loss of signals required for the maintenance of the tailbud progenitor population, and without

these signals cells of the tailbud either undergo a cell fate change, or are removed through apoptosis. This is most likely due to misexpression of two pro-neural genes, *Irx1* and *Irx2*. We provide evidence that their misexpression is most likely due to a mutation in a shared regulatory region.

Materials and methods

Animals

Fertilized chicken eggs were obtained from SkyBlueEgg (Arkansas, U.S.A.) and the Clemson University Poultry Farm. Eggs were incubated at 38.5 °C in a humidified chamber to the desired stage. Embryos were staged according to Hamburger and Hamilton [31]. Skeletal material was the gift of the Araucana Club of America.

Bone and cartilage staining

Bone and cartilage staining was carried out on E18 Araucana^{Rp} and tailed controls using Alcian blue (Polysciences) and Alizarin red S (Acros Organics) according to standard procedures [32]. Briefly, Embryos were fixed 3 x 24 hours in 95% EtOH, 100% EtOH, 2 x 24 h in 100% Acetone. Cartilage staining (20 mg Alcian Blue in 100 ml of 40% acetic acid glacial/EtOH) was performed from a few hours to overnight depending on sample size. Embryos were rinsed in EtOH for 15 min followed by EtOH for 24 hrs. They were then placed in saturated borax solution 2 x 24 hours ($Na_2B_4O_7 \cdot 10H_2O$ in H_2O). Trypsin solution (0.45g purified trypsin in 400mL of 30% borax dissolved in distilled water) at 30°C was used to clear tissue until flesh became translucent and soft (between 1-4 days, depending on size of sample). Alizarin Red S solution (0.5% KOH and 0.1% Alizarin Red S) was used to stain bones (12-24 hours). Samples were then washed in distilled water, followed by a wash in 0.5% KOH solution for 15 min. Excess Alizarin Red S stain was removed using 0.5% KOH solution for 2 x 24 hours at room temperature under a light source. Samples then went through series of glycerol 0.5% KOH washes (20%

glycerol/0.5% KOH, 50/50 and 75/25 mix). Samples were stored in 100% glycerol with 100mg Thymol crystals.

Somite number counts

Araucana^{Rp} and controls were incubated to between HH16-25. Embryos were harvested and somite counts performed using a Nikon stereoscopic microscope. At later stages, between HH22-25, *Dact2* ISH labeling was used to aid counts of the posterior somites.

Statistical Analysis

Assuming a normal distribution of the data, a two-tailed *t*-test was carried out to test for differences in the average values of samples from experiments for somite counts, proliferation, TUNEl, and flow cytometry. Analysis was carried out using Statistical Analysis Software (SAS).

Immunohistochemistry

Embryos were fixed in 4% paraformaldehyde (PFA/PBS) for 48 hours before being cryoembedded in 15% sucrose/7.5% gelatin/PBS and sectioned on a Leica cryotome at 25 μm. Immunostaining was carried out using our standard protocol [33]. Briefly, sections were blocked in PBS with 0.1 % TritonX-100 and 0.2% bovine serum albumin (BSA). Then incubated overnight at 4 °C with the primary antibodies anti-E-cadherin (cat 610182, BD Bioscience) and anti-laminin (cat L9393, Sigma). Following washing in PBS sections were incubated at secondary antibodies 1:200 Alexa Fluor 488 goat anti-mouse

IgG and Alexa Fluor 594 goat anti-rabbit IgG (Invitrogen). Following washing in PBS and mounting with SlowFade (life technologies), fluorescent images were captured using a Nikon Ti Eclipse confocal microscope.

In situ hybridization

Whole-mount in situ hybridization was performed according to our standard procedures using probes against *Brachyury*, *Cyp26a1*, *Dact2*, *Fgf8*, *Irx1*/2/4 *Meso1*, *Raldh2*, *Sox2* *Tbx6*, *Wnt3a*, [33]. The probes have all been previously described as follows: *Dact2*, *Meso1* and *Tbx6*, [6], *Fgf8* [34], *Irx1*, *Irx2*, *Irx4* [35], *Raldh2* [36] and *Wnt3a* [37]. *Lfng* probe was the generous gift of Olivier Pourquié [38]. *Irx1 Irx2* and *Irx4* probes were the generous gift of Dr. Cheryll Tickle (*Irx1* Accession# NM_001030338.1, base pairs 1005-1404 and *Irx2* Accession# NM_001030336.1, base pairs 1227-1434). Embryos were cryoembedded in 15% sucrose/7.5 % gelatin/PBS and sectioned on a Leica cryotome. Whole mount embryos and sections were imaged on a Nikon Smz1500 stereomicroscope and Nikon Eclipse 80i compound microscope, respectively using a Qimaging Micropublisher 5.0 camera.

EdU and TUNEL labeling

For proliferation analysis, Click-iT EdU 488 Imaging Kit (Invitrogen) was used to carry out labeling of cells as previously described [39]. Briefly, embryos were pulsed with EdU for 60 minutes before harvesting and fixation in 4% PFA overnight. Embryos were then cryoembedded in 15%sucrose/7.5%gelatin/PBS and sectioned on a Leica cryotome at

25µm. Alternating sections were processed for EdU detection, or apoptosis detection using the TUNEL method (In Situ Cell Death Detection Kit, TMR red, Roche), and imaged on a Nikon Ti Eclipse confocal microscope. Image analysis of EdU and TUNEL labeling was carried out using NIS Elements (build 736) cell counting software.
EdU labeling for flow cytometry analysis was carried out for 60 minutes. Following harvesting, cells were dissociated and fixed as previously described [40]. EdU detection was carried out following the manufacturers instructions. Cells were filtered through a 60µm mesh to make a single cell suspension. Flow cytometry was run on a Guava easyCyte. Analysis was carried out using CytoSoft5.3 software.

Whole genome sequencing and bioinformatics

DNA samples for six Araucana were acquired from our previous study [30]. Each of the six samples was sequenced on six lanes with an Illumina HiSeq 2000 sequencer. Average genomic coverage was 27.63x and average number of bases sequenced was 29.009 Giga base pairs (Table S1). Sequence reads were trimmed using Trimmomatic and aligned to the corresponding region previously identified to be associated with the rumpless phenotype on chromosome 2 of the ICGSC Gallus_gallus-4.0/galGal4 build using Bowtie2 applications [41] [30] [42]. The mpileup function of SamTools was used to call variants [43]. The view option of bcftools was used to call the genotype at each variant for each individual bird using the bcftools defaults. Variants that were found to be homozygous in all three homozygous rumpless birds, found to be heterozygous in all three heterozygous birds, and were not found in the homozygous tailed bird were

considered fixed in the population and targets for genomic variation that may result in the rumpless phenotype. Identified variants were compared against known variants in the Beijing Genomics Institute (BGI) database, and variants that were previously identified not to be involved in rumplessness were removed [44]. Variants were also compared to previously identified highly conserved non-coding regions [45].

Results

Araucana^{Rp} lack the caudal-most vertebrae

The current North American breed standard of the Araucana chicken requires that they must lack the tail region, called rumplessness (Figure 1A). A comparison of adult tailed control and Araucana^{Rp} skeletons show this distinctive morphology. Rumplessness is due to the loss of the free caudal vertebrae and pygostyle of the tail (Figure 1B-C). The severity of the phenotype changes slightly among different animals, thus we found that in some cases Araucana^{Rp} skeleton have 1-3 caudal vertebrae. We performed cartilage/skeletal staining to determine if the lack of these vertebrae is due to a lack of formation, or if they form and are reabsorbed. To that end, embryos were processed for Alcian Blue and Alizarin Red S to label cartilage and bone, respectively. By embryonic day (E) 18, control embryos have formed all of the 11 caudal free vertebrae and the pygostyle (Figure 1D). However, Araucana^{Rp} embryos lack the free caudal vertebrae and the pygostyle (Figure 1E). From these data we conclude that the rumpless phenotype observed in the Araucana^{Rp} adult chicken appears during development because of the lack of formation of the caudal vertebrae and the pygostyle. These data match previous observations with rumpless chickens [27].

Araucana^{Rp} embryos display truncated tail morphology and down regulation of tailbud signals

We next investigated the stage at which the precursors to the caudal vertebrae form. To that end we collected control and Araucana^{Rp} embryos during somitogenesis. Until HH

14-15 we did not observe any gross morphological differences. Beginning at HH16 the embryonic tailbud is visibly changed in morphology in Araucana^{Rp} embryos compared to controls (Figure 2 A,D). In Araucana^{Rp} embryos the tail appears shorter and more pointed, with the angle of curvature being wider than in controls (Figure 2A,D). As tail extension continues at HH18, Araucana^{Rp} embryo tails do not extend as far as controls (Figure 2B,E). By HH20 the Araucana^{Rp} embryo tail is truncated compared to controls (Figure 2C,F). The most recently formed somite (asterisk) is much closer to the tip of the tail in Araucana^{Rp} compared to controls (Figure 2C,F), suggesting that the size of the non-segmented paraxial mesoderm, which is the precursor of the somites, is being drastically reduced. We then investigated these anomalies using molecular markers.

Brachyury is a transcription factor expressed in the tailbud mesenchymal region that is required for mesoderm and notochord formation. Mutations in *Brachyury* are associated with problems in tail growth and the formation of the sacral vertebrae of the mouse embryo [22]. After staining the embryos by *in situ* hybridization (ISH) we found that the expression at HH15 and earlier stages is unchanged between Araucana^{Rp} and control embryos (data not shown). However, beginning at HH16, *Brachyury* is down regulated within cells of the tailbud region of the Araucana^{Rp} embryos (Figure 2, G-L). Expression remains unaffected within the notochord (Figure 2J-L). This result indicates that the population of progenitor cells responsible to generate the paraxial mesoderm is affected in the Araucana^{Rp} embryos. Next we investigated the expression of *Tbx6*, which marks the non-segmented paraxial mesoderm from which somites are formed (Figure 2M-O).

No differences are seen between controls and Araucana^{Rp} earlier than HH15 (25 somites) (data not shown). Starting at HH15 (25 somites), *Tbx6* expression starts to be down regulated in Araucana^{Rp} tailbud (Figure 2P and inset). Expression is lost within the population of cells posterior to the forming neural tube and notochord as well as the cells of the tailbud mesenchyme. This region corresponds to the most recent group of cells leaving the progenitor population in the tailbud, suggesting that in Araucana^{Rp} embryos the newly formed tissue is not acquiring a paraxial mesoderm fate.

Expression of *Tbx6* is maintained in the more anterior cells of the PSM, with the domain of expression being reduced in size until HH20, probably due to the normal periodic incorporation of the most anterior cells into the newly formed somites. These results show that the defect leading to Araucana^{Rp} rumplessness arises early in tailbud development, and involves the down regulation of signals required for specification of the paraxial mesoderm. As early as HH15 the production of new paraxial mesoderm in the Araucana^{Rp} embryo is arrested.

Somitogenesis ends prematurely in Araucana^{Rp} embryos

The failure of full tail elongation during development, along with the down regulation of signals required for specification of the paraxial mesoderm suggests a failure of paraxial mesoderm specification, including a failure of somitogenesis in Araucana^{Rp} embryos. In normal conditions, tail somitogenesis continues until HH24-25, when all 51-53 somites have formed [6].

In order to understand when somitogenesis stops in Araucana^{Rp}, we carried out ISH labeling to determine the precise number of somites formed. To that end we tested the expression of *Meso1*, which is important for the delineation of new somites. *Meso1* is expressed in the anterior most presomitic mesoderm, marking an area posterior to the next pair of forming somites and encoding a transcription factor required for the formation of the new epithelial somite [46][47] [46]. In normal chick embryos *Meso1* expression is not down regulated until HH24-25, which marks the end of somitogenesis [6]. We found that *Meso1* expression is unchanged in Araucana^{Rp} embryos through HH18 compared to controls (Figure 3A-B and D-E). At HH19, control *Meso1* expression remains up regulated, (Figure 3C) whereas in Araucana^{Rp} *Meso1* expression is lost (Figure 3F). This result indicates that somite formation is arrested at HH19 rather than at HH24-25.

We also tested the expression of *Dact2*, a marker of the forming somite and most recently formed somites [47] [6]. We found that at HH18-19 *Dact2* expression in the Araucana^{Rp} embryos is similar in the number of recently formed somites when compared to controls, but the distance from the most posterior labeled somite to the tip of the tail appears shorter (Figure 3G-H and J-K). By HH20 *Dact2* expression in Araucana^{Rp} labels fewer forming somites and is down regulated compared to controls (Figure 3I and L).

The down regulation of *Meso1* and *Dact2* as much as 6 stages earlier than is seen in tailed controls together with the premature loss of the paraxial mesoderm marker *Tbx6*

suggests a failure of formation of the most posterior somites. To test this idea we compared the total number of somites in Araucana^{Rp} and controls embryos between HH stages 16-25 (Araucana^{Rp} n=83, control n=73, Figure 3M). The number of control somites at each stage closely matches the expected number of somites as described in the normal stage series [31]. Beginning at HH stage19, the number of somites in Araucana^{Rp} embryos is significantly different than controls, and by the end of somitogenesis at HH25, Araucana^{Rp} embryos form on average 11 fewer somites than control embryos (Figure 3M). These data are consistent with our bone and cartilage staining results in which embryos can lack the 5 free caudal vertebrae and the 6 vertebrae of the mature pygostyle. In summary, these data show that somitogenesis ends prematurely within the tailbud of Araucana^{Rp} embryos, and that new somites fail to form as early as HH20, once all the previously specified paraxial mesoderm has been incorporated into the somites.

To investigate if the failure to form all 51-53 somites in Araucana^{Rp} is due to a slower rate of somite formation we analyzed the speed of somite formation, similar to what has been previously performed [6]. To that end, control and Araucana^{Rp} embryos with a given number of somites were cultured for fixed periods of time, 18-24 hours, and counted again to determine the number of additional somites that were made. This analysis showed no differences between controls and Araucana^{Rp} embryos, which indicates that speed of somite formation is not affected in Araucana and cannot explain the rumples phenotype (data not shown). Similarly, we checked cyclic gene expression by ISH using a *Lunatic fringe* probe to detect its expression in the PSM. *Lunatic fringe*

has a cyclical period of expression equal to the time of somite formation, which in the chick embryo is 90 minutes [6] [38]. The analysis revealed no differences between controls and Araucana^{Rp} embryos (data not shown).

Ingression through the VER is unchanged in Araucana^{Rp}

As the morphological defects in the Araucana^{Rp} embryo are observed at the same stages in which the embryo undergoes the transition from the primary body formation, which is based in the primitive streak and Hensen's node, to that of secondary body formation, which is based in the tailbud, suggests that these two events could be related. To that end we analyzed the expression of markers of components involved in this transition. As mentioned earlier, cells undergo an epithelial to mesenchymal transition, ingressing through the VER to contribute to the tailbud mesenchyme, which is later followed by differentiation into definitive lineages [17] [4]. In normal embryos, the process is arrested at later stages when the tailbud territory is exposed to noggin, which then inhibits BMP signaling responsible for the epithelial to mesenchymal transition and basal membrane degradation [4]. If basal membrane breakdown is stopped prematurely then ingression of cells into the tail is blocked leading to a truncated tail phenotype [4]. We examined this process by labeling the VER of the control and Araucana^{Rp} embryos with Laminin, a marker of the basal membrane, and E-cadherin, a cell-cell adhesion marker. We found that at HH18 breakdown of the basal membrane is occurring in both control and Araucana tailbud (Figure S1). These results show that the rumpless morphology of the

Araucana^{Rp} embryo is not due to morphological problems in the progenitor regions associated to the transition from primary to secondary body formation.

Sequencing of the critical region reveals candidate mutations within a known regulatory region.

The causative mutation is likely one of two possible types for the observed Araucana^{Rp} phenotype. Either the coding region of one or both genes is affected, or there is a change in regulation of these genes [30]. To analyze the 740kb critical region we performed whole genome sequencing on DNA from six Araucana birds. After aligning reads to chromosome 2, variants (insertions, deletions, and SNPs) were called using the mpileup function of SamTools. Variants were separated into three haplotype groups: three homozygous rumpless, two heterozygous, and one homozygous tailed Araucana.

A total of 2092 unique small variants were identified within the candidate region when compared to the chicken gal gal4.0 reference sequence. We reduced the list of variants by excluding those that did not hold true to type, keeping only those that occurred in every homozygous, half of heterozygous and never in tailed Araucana. A total of 316 small variants matched this pattern in the rumpless region. A further 18 variants that lined up with previously reported variants in tailed birds were also removed [44]. Of the remaining 298 variants, we identified 274 SNPs and 24 insertion/deletions unique to rumpless Araucana (Figure 4). None of the identified 298 small variants fell within the exons or introns of *Irx1* or *Irx2*. These results support the idea that the candidate mutation

lies in the intergenic or flanking regions of *Irx1* and *Irx2*. Comparison of the 298 candidate variants with highly conserved non-coding regions within the candidate region revealed a subset of candidate variants falling within 4 of the highly conserved non-coding regions (Table S2) [45]. PCR analysis of these regions is underway in a large cohort birds from multiple breeds.

Irx1 and *Irx2* are misexpressed in the Araucana^{Rp} embryo tailbud

Iroquois-class homeodomain proteins play multiple roles during pattern formation of vertebrate embryos specifying the identity of diverse territories of the body. One of their primary roles is the initial specification of the vertebrate neurectoderm [48] [49]. Importantly, *Irx1* and *Irx2* are not expressed in the tailbud territory of the normal chick embryo at any stage of development. To investigate if indeed the Araucana^{Rp} embryo displays an aberrant expression of *Irx1* and *Irx2* we determined their pattern of expression around the stages in which the defects associated to rumplessness start. Expression of *Irx1* in control embryos is restricted to the neural tube of the elongating posterior axis (Figure 5A-C). Analysis of expression of *Irx1* before HH15 revealed no change in expression between controls and Araucana^{Rp} (Figure 5A,D). However, at HH15 Araucana^{Rp} tailbud misexpressed *Irx1* with a restricted pattern of expression similar to that observed for the region containing the progenitor population (Figure 5E). This misexpression of *Irx1* is the earliest change in gene expression we have identified in Araucana^{Rp} embryos. Anterior near the level of recently formed somites, expression of *Irx1* in the neural tube is normal (Figure 5M). A combination of sagittal and transverse

sections shows that *Irx1* misexpression in Araucana^{Rp} is at the level of the chordoneural hinge at HH15 (Figure 5N-P)**.** *Irx1* misexpression in Araucana^{Rp} tailbud is maintained at HH16 (Figure 5F) and by HH18 the expression disappears (data not shown). Similarly, *Irx2* is misexpressed in Araucana^{Rp} tailbud region beginning at HH15 until it is down regulated by HH17 (Figure 5K; data not shown). Transverse sections clearly show that misexpression of *Irx2* is located at a similar posterior level as the misexpression of *Irx1*, although located slightly more lateral (Figure 5Q).

Although *Irx4* does not fall within the defined critical region, it is possible that its expression could be altered along with *Irx1* and *Irx2* because the three genes are part of the same genomic cluster [45]. ISH analysis of *Irx4* expression revealed no differences between control and Araucana^{Rp} embryos (data not shown). Thus, despite their proximity in a genomic cluster, the aberrant expression of *Irx1* and *Irx2* found in the Araucana^{Rp} tailbud is not affecting all Iroquois-class genes, even from the same cluster. The misexpression of *Irx1* and *Irx2* in the Araucana^{Rp} tailbud combined with these two genes being the only genes within the critical genomic region highly suggests that the Araucana^{Rp} phenotype is due to a change in regulation of these genes [30].

Loss of maintenance of the tailbud progenitor population in Araucana^{Rp} embryos

The *Iroquois* genes are known to be involved in specifying and patterning neural domains [48] [49]. In *Xenopus*, over-expression of Iroquois genes leads to neural plate expansion, and the onset of neural differentiation is advanced [50] [51]. Cells of the tailbud have a

bipotential fate, and can become mesoderm (paraxial mesoderm) or ectoderm (neural tube) [12]. If exposure of the CNH to a pro-neural signal in the form of *Irx1* and *Irx2* pushes cell fates towards neural, we would predict to see a down regulation of signals associated with the tailbud mesoderm and maintenance of the tailbud progenitor population, and the up regulation of proneural markers. To that end, we determined the expression of markers necessary for continued elongation and maintenance of the tailbud progenitor population, such as *Wnt3a* and *Fgf8* [19][54].

Wnt3a expression is limited to the tailbud and dorsal neural tube. We found that until HH16 *Wnt3a* is indistinguishable between control and Araucana^{Rp} tailbuds (Figure 6A, D). On the other hand, by HH17 *Wnt3a* expression becomes significantly down regulated and is lost by HH18 (Fig 6B-C, E-F).

We next tested the expression of *Fgf8*, which in control embryos is expressed within the tailbud through tail elongation (Figure 6G-I). We found that prior to HH17 expression of *Fgf8* in Araucana^{Rp} tailbud is unchanged from controls (data not shown). However, the expression is clearly down regulated as early as HH17 (Figure 6J) and by HH18 *Fgf8* is completely lost (Figure 6K-L). These data indicate that in the Araucana^{Rp} tailbud and following the initial aberrant expression of *Irx1/Irx2* there is a loss of the signals responsible for maintaining the undifferentiated progenitor population.

Fgf8 and Wnt3a work together to maintain the tailbud progenitor population in an undifferentiated state [52]. This function can be disrupted by exposure to RA, which appears to be one of the mechanisms invovled in the physiological termination of the process of axis elongation [6]. In addition, when a control embryo is exposure to ectopic RA it eliminates *Wnt3a* and *Fgf8* from the tailbud, and blocks the axial growth of the embryo [6]. Similarly, in *Cyp26a1*-/- embryos, which lack the expression of an enzyme that degrades RA produced in the somites [23], the progenitor population is unprotected against RA and the result is the generation of embryos suffering severe caudal truncation [23]. It is therefore possible that the loss of *Wnt3a* and *Fgf8* in the Araucana^{Rp} tailbud could be due to a non-regulated exposure to RA. Thus, we decided to investigate the possible implication of RA in the process.

First, we evaluated the expression of *Raldh2*, which encodes a dehydrogenase involved in endogenous production of RA (Figure 6S-U). *Raldh2* expression in the rostral end of the PSM and the newly formed somites continues to advance posteriorly as the tailbud elongates. We found that expression of *Raldh2* in Araucana^{Rp} somites appears normal compared to the pattern displayed in control embryos (Figure 6V-X). . We then evaluated the expression of *Cyp26a1* and found that at earlier stages there is no difference in expression in the tailbud of Araucana^{Rp} and control embryos (data not shown). However, from HH16 *Cyp26a1* is completely lost in the Araucana^{Rp} tailbud (Figure 6M-R). In summary, these results show that following misexpression of *Irx1* and *Irx2* in Araucana^{Rp}, there is a down regulation of *Cyp26a1,* leaving the tailbud unprotected and

exposed to the RA gradient originating from the somites. After that, *Wnt3a* and *Fgf8* are both down regulated within the tailbud, suggesting the progenitor cells of the tailbud may be undergoing premature differentiation and/or apoptosis due to RA exposure.

Expansion of neural tissue within Araucana^{Rp} tailbud

Considering the combination of ectopic expression of *Irx1* and *Irx2* (both pro-neural) and the loss of protection from the effects of RA due to the down regulation of *Cyp26a1*, we predicted that most of the tailbud cells would be pushed towards a neural fate. To investigate this possibility we tested the expression of *Sox2*, which is expressed in the forming neural cells [53]. We found the more anterior neural tube is normal in the Araucana^{Rp} embryos and consists of a single lumen that expresses *Sox2* (data not shown). Beginning at HH18, Araucana^{Rp} embryos have ectopic neural tissue expressing *Sox2* with multiple irregular lumens within the tail region, compared to single lumen seen in controls (Figure A-J). Thus, as expected, the Araucana^{Rp} neural tube appears to have an enlarged epithelial structure with multiple open and irregular lumens (Figure 7 J).

Further evidence for ectopic neural tissue at later stages can be seen with *Wnt3a* expression in the dorsal neural tube. Control expression of *Wnt3a* is limited to the dorsal neural tube (Figure 7 K, L), whereas Araucana^{Rp} expression of *Wnt3a* is dramatically altered in the tailbud region (Figure 7 M, N). This expression appears uncoordinated with similarly staged embryos each displaying different patterns (Figure 7 M, N). Thus, our results demonstrate the aberrant neural differentiation of the progenitor population in the

Araucana^{Rp} tailbud, which is probably one of the main reasons for the generation of the rumplessness phenotype.

Araucana^{Rp} tailbud displays decreased proliferation and increased apoptosis

The premature down regulation of markers associated to the proliferation of the progenitor population, such as *Fgf8* and *Wnt3a*, together with the extended domain of expression of the neural marker *Sox2* indicates that the progenitor population has been critically jeopardized by their forced differentiation towards neural cell fate. To further support this idea we evaluated proliferation in the Araucana^{Rp} tailbud. To that end we performed Edu labeling. Until HH16 we did not find any difference between control and Araucana^{Rp} samples, however beginning at HH17 there is an observable decrease in the number of proliferating cells within the Araucana^{Rp} tailbud mesenchyme region containing the progenitor population (Figure 8 C-D, G-H). However, quantification by cell count of the number of proliferating cells within the tail reveals no statistical difference in the ratio of proliferating/total cells between control and Araucana^{Rp} tails (Figure 8I). Thus, although there appears a reduction of proliferation consistent with the observed down regulation of the markers involved in the maintaining of the progenitor population *Wnt3a* and *Fgf8*, it is not significantly different than the level of proliferation observed in control tails.

Exposure to high levels of RA has been shown to lead to apoptosis [54] [55]. We tested if there is increased apoptosis within the tailbud region. Thus, we performed TUNEL

staining on sagittal sections of the tail region to label dying cells. No difference was seen at or before HH16 between control and Araucana^{Rp} embryos (Figure 8 J-K, N-O). Staining revealed that beginning at HH17 there is a clear and significant increase in apoptotic cells in Araucana^{Rp} compared to controls (Figure 8 L-M, P-Q). Apoptosis was primarily localized to the posterior most cells of the tailbud, the region known to contain the progenitor population (McGrew et al., 2008). Quantification of TUNEL positive cells revealed that beginning at HH17-18 there is a statistically significant increase in apoptosis in Araucana^{Rp} compared to controls (Figure 8 R). This indicates that, in addition to the neural transformation, some cells from the progenitor population of the Araucana^{Rp} tailbud are dying by apoptosis.

Discussion

To our knowledge, the Araucana^{Rp} is the first reported example in which aberrant expression within the tailbud of a developing vertebrate embryo seems to be responsible for the malformations associated to the process of segmentation and axial. We have shown that the Araucana^{Rp} phenotype results from a failure to form the caudal most vertebrae due to the lack of formation of the last somites. Sequencing of the candidate rumpless region revealed a number of small variants centered in an area containing the pro-neural genes *Irx1* and *Irx2*. We found that both genes are misexpressed within the tailbud region containing the progenitor population responsible of the axial growth of the Araucana^{Rp} embryo. After this misexpression, there is an immediate down regulation of *Cyp26A1*, potentially allowing RA secreted from the somites to target the tailbud. This is

quickly followed by down regulation of *Wnt3a* and *Fgf8,* which are required for maintenance and proliferation of the tailbud progenitor population. The down regulation of *Brachyury* and *Tbx6* indicates there is an arrest in the production of additional paraxial mesoderm necessary for the generation of the somites, whereas the increase in *Sox2* expressing tissue within the tail, and the presence of ectopic neural tubes are consistent with the premature differentiation of the progenitor cells towards neural cell fate. Last, the exposure to RA is probably responsible for the increased apoptosis that affects the tailbud progenitor population.

Araucana^{Rp} embryos fail to form the caudal-most somites, in the more severe cases missing the last 11 caudal somites. Somite formation stops around HH20, 4-5 stages earlier than control embryos, as seen by the down regulation of *Meso1* and *Dact2* . This observation is in agreement with a model of failure to form somites proposed 70 years ago by Zwilling (1942) versus the alternative possibility in which somites are removed or reabsorbed [27]. Morphological changes in the shape of the tail are evident as early as HH16 and coincide with the down regulation of mesodermal tailbud markers, *Tbx6* and *Brachyury*, the loss of which have been documented to lead to truncation of the tail region [56] [20]. Araucana^{Rp} embryos display some variability in the final number of somites, and this appears to equate to variation in the number of caudal vertebrae as seen by cartilage and bone staining at E9-E18.

Araucana^{Rp} skeletons also show a range of phenotypes, some having 1-2 free caudal vertebrae, the caudal-most appearing fused, whereas other bones exhibit no appearance of any free caudal vertebrae. These variations seem to be inherent to the Araucana breed and were previously described in the initial reports on the rumpless phenotype [57] [58]. The complete lack of free caudal vertebrae was described as true rumplessness, whereas the presence of 1-3 free caudal vertebrae was described as intermediate rumpless. However, intermediate birds were found to carry the rumpless gene as evidenced by breeding experiments, thus it remains unclear how the intermediate phenotype arises [57]. We have previously reported that partial tailed Araucana are heterozygous for the rumpless linkage region [30]. One possibility is that intermediate rumpless Araucana may be due to changes in expressivity, possibly owing to other modifiers outside of the previously identified linkage.

As of the current galGal4 assembly of the chicken genome created by the University of California, Santa Cruz (http://genome.ucsc.edu/), *Irx1* and *Irx2* are the only two genes contained within the 0.74 Mb critical region linked with the rumpless phenotype [30]. We have identified 298 candidate small variants within the rumpless critical region. None of the candidate mutations are located within the coding sequence of *Irx1* or *Irx2* or their intronic regions. It would be expected that any changes in coding sequence of *Irx1* and/or *Irx2* would also lead to changes in patterning of the brain, where both genes are normally expressed. But in fact, Araucana^{Rp} embryos do not display differences outside of the tail. Tena and colleagues have found that the *Irx1* and *Irx2* promoters located near the two

genes are brought close together through a 3D architecture, allowing them to share enhancers [45]. Using these shared regulatory elements *Irx1* and *Irx2* are co-regulated. Based on the dual misexpression of *Irx1* and *Irx2* in Araucana^{Rp}, it is highly probable that the causative mutation lies within a regulatory element, thus changing the regulation of both genes. The causative mutation could act by removing specific repression/silencing of *Irx1* and *Irx2* within the tailbud, or by creating a new enhancer, specific to the tailbud. Although we were able to identify a number of mutations within predicted regulatory regions, a more detailed analysis of this aberrant activation will allow us to determine the precise mechanism involved.

From HH15 both *Irx1* and *Irx2* are misexpressed within the tailbud region of the Araucana^{RP} embryos preceding all other changes we have been able to identify in gene expression and morphology related to the rumpless phenotype. These two *Iroquois* genes encode the TALE (three amino acid loop extension) family homeodomains and are involved in proneural patterning [59] [60] [61] [51]. During development of a normal chick embryo, expression of *Irx1* and *Irx2* is limited to neural tissue: the developing brain and neural tube. It was therefore expected that the ectopic expression of *Irx1/Irx2* would generate an expansion of the neural territory. That is precisely what we found, as shown by the abnormal enlargement of the *Sox2* expression domain. Further analysis will be required to determine if this neural induction is indeed a direct effect of *Irx1* and/or *Irx2* expression. It is very surprising that the misexpression of *Irx1* and *Irx2* in the tailbud is restricted to a very specific window of development. The expression starts at HH15 and

disappears again at HH17. This temporal and special restriction suggests the implication of a mechanism or factor with similar restrictions.

During somitogenesis, the caudal region of the vertebrate embryo including the PSM is exposed to RA expressed from the somites and Fgf8 expressed from the tailbud, generating two opposing gradients [62][55][56]. The tailbud progenitor population must be protected from the RA secreted from the somites by the action of *Cyp26a1*, which metabolizes RA [23]. It is indicative that one of the first consequences of the misexpression of *Irx1*/*Irx2* is the loss of *Cyp26a1* within the tailbud. However, whether the down regulation of *Cyp26a1* is a direct consequence of the ectopic *Irx1/Irx2* expression is still unknown. Data produced using chicken and mice embryos show that ectopic RA leads to the down regulation of *Wnt3a* and *Fgf8*, axis truncation, and ectopic neural tissue [54] [6]. In the tailbud of Araucana^{Rp}, without the protection of *Cyp26a1*, RA can target the territory, which would induce the down regulation of *Wnt3a and Fgf8*, as we observed [63] [6] [55]. Work in mice has shown that loss of *Wnt3a* signaling leads to axis truncation [64]. In addition, in these *Wnt3a* mutants there is ectopic neural tissue formed, suggesting that *Wnt3a* is required for proper patterning of presumptive paraxial mesoderm cells, and without *Wnt3a* there is an expansion of neural tissue [65] [66]. Down regulation of *Wnt3a* in Araucana^{Rp} then, would mean a further loss of the ability to specify paraxial mesoderm.

In fact, following the misexpression of *Irx1* and *Irx2*, the mesodermal patterning transcription factors *Tbx6* and *Brachyury* are down regulated. It is possible that their down regulation is a consequence of the parallel elimination of *Wnt3a* and *Fgf8*, as the expression of all these genes is inter-related [20,67] [18]. *Brachyury* is required for proper mesodermal cell fate and without its expression embryos fail to form the proper mesodermal structures [68]. Thus, *Brachyury*(-/-) embryos fail to form both somites and the notochord, with heterozygotes forming a truncated axis [21] [69]. In Araucana^{Rp} embryos *Brachyury* is down regulated within the posterior tailbud population, whereas expression in the notochord is maintained, which indicates that the effect of *Irx1*/*Irx2* misexpression is restricted to the cells of the tailbud. Interestingly, the observed down regulation of *Brachyury* only within the tailbud cells closely resembles that found in zebrafish embryo tails, whose expression of *no-tail* (*ntl*) (zebrafish *Brachyury* ortholog) is down regulated upon treatment with RA or through down regulation of *Wnt3a* [70] [71].

The absence of *Tbx6* in the homozygous mutant mouse embryo leads to a loss of paraxial mesoderm specification, and an up regulation of the neural marker *Sox2* within ectopic neural tubes [56] [65]. Tbx6 indirectly regulates *Sox2* through repression of its N1 enhancer, making it both necessary to repress neural fate as well as push cells towards a paraxial mesoderm fate [72]. *Tbx6* is a very specific marker for the paraxial mesoderm tissue and its expression is initiated just at the moment the progenitor cells divide and differentiate acquiring the paraxial mesoderm fate. In Araucana^{Rp}, similar to *Brachyury*,

Tbx6 expression is at first only lost within the posterior tailbud, not affecting *Tbx6*-positive cells positioned more anterior, which were specified previously. At the anterior end of the PSM, the formation of the somites continues unaffected, as indicated in our analysis, consuming periodically the PSM tissue until the process reaches the last *Tbx6*-positive cells. And when these remaining *Tbx6*-cells are incorporated in the last somite, the somitogenesis is prematurely arrested. Our data suggest a loss of signaling within the posterior tailbud that is required for proper mesodermal patterning. Without the proper signals, the presomitic paraxial mesoderm (labeled with *Tbx6*) begins to shorten. Without Tbx6 to both repress a neural fate as well as promote a paraxial mesoderm fate, along with Brachyury, cells that would normally form paraxial mesoderm form ectopic neural tissue, or remain in an undifferentiated transition state.

The lack of change in apoptosis before HH17 indicates that apoptosis is not the main cause of the morphological malformations associated with the Araucana^{Rp} tailbud. It is only after *Cyp26a1*/*Tbx6* has been lost that there is an observable increase in apoptosis, primarily centered on the tailbud region containing the progenitor population. Interestingly, this pattern appears very similar to the physiological apoptosis observed at the end of somitogenesis at HH25 in the chicken tailbud, which has been proposed to remove the remaining unsegmented PSM and progenitor population [55] [6] [5]. Similarly, the lack of change of proliferation within the tailbud prior to HH17 indicates that it is most likely not involved in the initial changes in tail morphology, but a consequence of the changes the territory is undergoing. Consistent with this idea is that

the observed decrease in proliferation within the tailbud primarily overlaps in distribution and time with the loss of those signals required for it, such as *Wnt3a* and *Fgf8*. However, as there is no statistical difference in the number of proliferating cells of the tail, it is difficult to say what role proliferation has in Araucana^{Rp}.

Interestingly, many of these processes, such as, down regulation of genes involved in tailbud progenitor maintenance such as *Fgf8*, *Wnt3a*, exposure to RA by down regulation of *Cyp26a1,* and an increase in apoptosis within the posterior tailbud population, are also observed during the normal end of somitogenesis. The specific mechanism RA reaches the tailbud is different in chick and mouse embryos. In the chick embryo there is an onset of *Raldh2* in the tailbud, which produces extra RA counteracting the protection given by *Cyp26a1*, whereas the mouse embryo loses *Cyp26a1* allowing the action of RA from the somites [6]. It is fascinating that the aberrant expression of *Irx1/Irx2* in the chicken tailbud appears to recapitulate precisely the situation described for the mouse embryo.

In summary, our results strongly suggest that the Araucana^{Rp} phenotype is due to a genetic mutation(s) that causes misexpression of two coregulated genes, *Irx1* and *Irx2*. Based on this, we propose a model in which misexpression of *Irx1/Irx2* is responsible of the down regulation of *Cyp26a1* in the tailbud. Without the protective activity of *Cyp26a1* the region is exposed to RA expressed in the somites. One of the first consequences of this is the down regulation of *Brachyury* and *Tbx6* followed by *Wnt3a* and *Fgf8*. The loss of these factors together with the action of RA induces the

differentiation of the progenitor population into neural fate, arresting the production of additional mesodermal cells. The process of somitogenesis continues normally at the anterior end consuming the remaining paraxial mesoderm until is extinguished. The second effect of the elevated RA is the induction of apoptosis among the remaining cells. It is unlikely that the Araucana^{RP} chicken is a unique situation as it is possible that a similar aberrant misexpression, *Irx1/Irx2* or any other gene capable of affecting the progenitor population, could be involved in the generation of malformations in other vertebrate species. In fact, we predict that similar mechanisms could be responsible for the rumpless or tailless breeds found in other animal models [73] [74]. It is very encouraging that our results could help to understand some of the human congenital pathologies in which the formation of axial structures is affected, and for which we still do not have an explanation.

Acknowledgements

We thank Ann Charles of Sky Blue Egg for supplying Araucana eggs, Fritz and Joyce Ludwig and Jocelyn Clarke for Araucana skeletons, as well as the Araucana Club of America for providing samples and assistance. We are grateful to Candler Paige, Rachel Bernstein, Ami Hughes, Amber Manley and Purvi Amin for technical assistance. We would like to thank the Clemson Light Imaging Facility for technical advice concerning microscopy, and Meg Staton of Clemson University Genomics Institute (CUGI) for bioinformatics analysis of the whole genome sequences. We would like to thank Cheryll Tickle, J. Kim Dale, and Olivier Pourquié for reagents. This study was supported by grants from Clemson University, the NIH (grant number DC009236), and INBRE (GM103499:13-2191).

Research reported in this publication was supported by NIDCD of the National Institutes of Health under award number R01DC009236.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Technical Contribution of the Clemson University Experiment Station.

This material is based upon work supported by the CSREES/USDA, under project number SC1700374.

References

1. Brand-Saberi B, Christ B (2000) Evolution and development of distinct cell lineages derived from somites. Curr Top Dev Biol 48: 1-42.

2. Gomez C, Ozbudak EM, Wunderlich J, Baumann D, Lewis J, et al. (2008) Control of segment number in vertebrate embryos. Nature 454: 335-339.

3. Richardson MK, Allen SP, Wright GM, Raynaud A, Hanken J (1998) Somite number and vertebrate evolution. Development 125: 151-160.

4. Ohta S, Suzuki K, Tachibana K, Tanaka H, Yamada G (2007) Cessation of gastrulation is mediated by suppression of epithelial-mesenchymal transition at the ventral ectodermal ridge. Development 134: 4315-4324.

5. Mills CL, Bellairs R (1989) Mitosis and cell death in the tail of the chick embryo. Anatomy and Embryology 180: 301-308.

6. Tenin G, Wright D, Ferjentsik Z, Bone R, McGrew M, et al. (2010) The chick somitogenesis oscillator is arrested before all paraxial mesoderm is segmented into somites. BMC Developmental Biology 10: 24.

7. Schoenwolf GC (1979) Histological and ultrastructural observations of tail bud formation in the chick embryo. The Anatomical record 193: 131-148.

8. Schoenwolf GC (1981) Morphogenetic processes involved in the remodeling of the tail region of the chick embryo. Anatomy and Embryology 162: 183-197.

9. Catala M, Teillet M-A, Le Douarin NM (1995) Organization and development of the tail bud analyzed with the quail-chick chimaera system. Mechanisms of Development 51: 51-65.

10. Cambray N, Wilson V (2002) Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. Development 129: 4855-4866.

11. Cambray N, Wilson V (2007) Two distinct sources for a population of maturing axial progenitors. Development 134: 2829-2840.

12. McGrew MJ, Sherman A, Lillico SG, Ellard FM, Radcliffe PA, et al. (2008) Localised axial progenitor cell populations in the avian tail bud are not committed to a posterior Hox identity. Development 135: 2289-2299.

13. Catala M, Teillet MA, De Robertis EM, Le Douarin ML (1996) A spinal cord fate map in the avian embryo: while regressing, Hensen's node lays down the

notochord and floor plate thus joining the spinal cord lateral walls. Development 122: 2599-2610.

14. Charrier JB, Teillet MA, Lapointe F, Le Douarin NM (1999) Defining subregions of Hensen's node essential for caudalward movement, midline development and cell survival. Development 126: 4771-4783.

15. Psychoyos D, Stern CD (1996) Fates and migratory routes of primitive streak cells in the chick embryo. Development 122: 1523-1534.

16. Freitas C, Rodrigues S, Charrier JB, Teillet MA, Palmeirim I (2001) Evidence for medial/lateral specification and positional information within the presomitic mesoderm. Development 128: 5139-5147.

17. Shook D, Keller R (2003) Mechanisms, mechanics and function of epithelial-mesenchymal transitions in early development. Mechanisms of Development 120: 1351-1383.

18. Wilson V, Olivera-Martinez I, Storey KG (2009) Stem cells, signals and vertebrate body axis extension. Development 136: 1591-1604.

19. Greco TL, Takada S, Newhouse MM, McMahon JA, McMahon AP, et al. (1996) Analysis of the vestigial tail mutation demonstrates that Wnt-3a gene dosage regulates mouse axial development. Genes & Development 10: 313-324.

20. Yamaguchi TP, Takada S, Yoshikawa Y, Wu N, McMahon AP (1999) T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. Genes & Development 13: 3185-3190.

21. Chesley P (1935) Development of the short-tailed mutant in the house mouse. Journal of Experimental Zoology 70: 429-459.

22. Herrmann BG, Labeit S, Poustka A, King TR, Lehrach H (1990) Cloning of the T gene required in mesoderm formation in the mouse. Nature 343: 617-622.

23. Abu-Abed S, Dollé P, Metzger D, Beckett B, Chambon P, et al. (2001) The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. Genes & Development 15: 226-240.

24. Ciruna B, Rossant J (2001) FGF Signaling Regulates Mesoderm Cell Fate Specification and Morphogenetic Movement at the Primitive Streak. Developmental Cell 1: 37-49.

25. Dunn LC, Landauer W (1934) The genetics of the rumpless fowl with evidence of a case of changing dominance. Journal of Genetics 29: 217-243.

26. Dunn LC, Landauer W (1936) Further data on genetic modification of rumplessness in the fowl. Journal of Genetics 33: 401-405.

27. Zwilling E (1942) The development of dominant rumplessness in chick embryos. Genetics 27: 641-656.

28. Landauer W (1945) Recessive rumplessness of fowl with kyphoscioliosis and supernumerary ribs. Genetics 30: 403-428.

29. Somes RG, Jr., Pabilonia MS (1981) Ear tuftedness: a lethal condition in the Araucana fowl. Journal of Heredity 72: 121-124.

30. Noorai RE, Freese NH, Wright LM, Chapman SC, Clark LA (2012) Genome-wide association mapping and identification of candidate genes for the rumpless and ear-tufted traits of the Araucana chicken. PLoS ONE 7: e40974.

31. Hamburger V, Hamilton LH (1951) A series of normal stages in the development of the chick embryo. Journal of Morphology 88: 49-92.

32. Yamazaki Y, Yuguchi M, Kubota S, Isokawa K (2011) Whole-mount bone and cartilage staining of chick embryos with minimal decalcification. Biotechnic & Histochemistry 86: 351-358.

33. Wood JL, Hughes AJ, Mercer KJ, Chapman SC (2010) Analysis of chick (Gallus gallus) middle ear columella formation. BMC Dev Biol 10: 16.

34. Chapman SC, Schubert FR, Schoenwolf GC, Lumsden A (2002) Analysis of Spatial and Temporal Gene Expression Patterns in Blastula and Gastrula Stage Chick Embryos. Developmental Biology 245: 187-199.

35. McDonald LA, Gerrelli D, Fok Y, Hurst LD, Tickle C (2010) Comparison of Iroquois gene expression in limbs/fins of vertebrate embryos. J Anat 216: 683-691.

36. Quinlan R, Gale E, Maden M, Graham A (2002) Deficits in the posterior pharyngeal endoderm in the absence of retinoids. Dev Dyn 225: 54-60.

37. Chapman SC, Brown R, Lees L, Schoenwolf GC, Lumsden A (2004) Expression analysis of chick Wnt and frizzled genes and selected inhibitors in early chick patterning. Dev Dyn 229: 668-676.

38. Dale JK, Maroto M, Dequeant ML, Malapert P, McGrew M, et al. (2003) Periodic notch inhibition by lunatic fringe underlies the chick segmentation clock. Nature 421: 275-278.

39. Warren M, Puskarczyk K, Chapman SC (2009) Chick embryo proliferation studies using EdU labeling. Developmental Dynamics 238: 944-949.

40. Hsieh YW, Yang XJ (2009) Dynamic Pax6 expression during the neurogenic cell cycle influences proliferation and cell fate choices of retinal progenitors. Neural Dev 4: 32.

41. Lohse M, Bolger AM, Nagel A, Fernie AR, Lunn JE, et al. (2012) RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Res 40: W622-627.

42. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9: 357-359.

43. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079.

44. Consortium ICPM (2004) A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. Nature 432: 717-722.

45. Tena JJ, Alonso ME, de la Calle-Mustienes E, Splinter E, de Laat W, et al. (2011) An evolutionarily conserved three-dimensional structure in the vertebrate Irx clusters facilitates enhancer sharing and coregulation. Nat Commun 2: 310.

46. Saga Y, Hata N, Koseki H, Taketo MM (1997) Mesp2: a novel mouse gene expressed in the presegmented mesoderm and essential for segmentation initiation. Genes Dev 11: 1827-1839.

47. Alvares LE, Winterbottom FL, Jorge EC, Rodrigues Sobreira D, Xavier-Neto J, et al. (2009) Chicken dapper genes are versatile markers for mesodermal tissues, embryonic muscle stem cells, neural crest cells, and neurogenic placodes. Dev Dyn 238: 1166-1178.

48. Cavodeassi F, Modolell J, Gomez-Skarmeta JL (2001) The Iroquois family of genes: from body building to neural patterning. Development 128: 2847-2855.

49. Gomez-Skarmeta JL, Modolell J (2002) Iroquois genes: genomic organization and function in vertebrate neural development. Curr Opin Genet Dev 12: 403-408.

50. Bellefroid EJ, Kobbe A, Gruss P, Pieler T, Gurdon JB, et al. (1998) Xiro3 encodes a Xenopus homolog of the Drosophila Iroquois genes and functions in neural specification. EMBO J 17: 191-203.

51. Gomez-Skarmeta JL, Glavic A, de la Calle-Mustienes E, Modolell J, Mayor R (1998) Xiro, a Xenopus homolog of the Drosophila Iroquois complex genes, controls development at the neural plate. EMBO J 17: 181-190.

52. Dequeant M-L, Pourquie O (2008) Segmental patterning of the vertebrate embryonic axis. Nat Rev Genet 9: 370-382.

53. Rex M, Orme A, Uwanogho D, Tointon K, Wigmore PM, et al. (1997) Dynamic expression of chicken Sox2 and Sox3 genes in ectoderm induced to form neural tissue. Dev Dyn 209: 323-332.

54. Shum ASW, Poon LLM, Tang WWT, Koide T, Chan BWH, et al. (1999) Retinoic acid induces down-regulation of Wnt-3a, apoptosis and diversion of tail bud cells to a neural fate in the mouse embryo. Mechanisms of Development 84: 17-30.

55. Olivera-Martinez I, Harada H, Halley PA, Storey KG (2012) Loss of FGF-Dependent Mesoderm Identity and Rise of Endogenous Retinoid Signalling Determine Cessation of Body Axis Elongation. PLoS Biol 10: e1001415.

56. Chapman DL, Papaioannou VE (1998) Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. Nature 391: 695-697.

57. Landaeur W (1928) The morphology of intermediate rumplessness in the fowl: With Remarks Concerning Hereditary and Accidental Rumplessness. Journal of Heredity 19: 453-467.

58. Landauer W, Dunn LC (1925) Two types of rumplessness in domestic fowls: A Morphological Comparison. Journal of Heredity 16: 153-160.

59. Gomez-Skarmeta JL, Diez del Corral R, de la Calle-Mustienes E, Ferre-Marco D, Modolell J (1996) Araucan and caupolican, two members of the novel iroquois complex, encode homeoproteins that control proneural and vein-forming genes. Cell 85: 95-105.

60. Rodriguez-Seguel E, Alarcon P, Gomez-Skarmeta JL (2009) The Xenopus Irx genes are essential for neural patterning and define the border between prethalamus and thalamus through mutual antagonism with the anterior repressors Fezf and Arx. Dev Biol 329: 258-268.

61. Gomez-Skarmeta J, de La Calle-Mustienes E, Modolell J (2001) The Wnt-activated Xiro1 gene encodes a repressor that is essential for neural development and down regulates Bmp4. Development 128: 551-560.

62. Dubrulle J, McGrew MJ, PourquiÈ O (2001) FGF Signaling Controls Somite Boundary Position and Regulates Segmentation Clock Control of Spatiotemporal Hox Gene Activation. Cell 106: 219-232.

63. Diez del Corral R, Olivera-Martinez I, Goriely A, Gale E, Maden M, et al. (2003) Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. Neuron 40: 65-79.

64. Takada S, Stark KL, Shea MJ, Vassileva G, McMahon JA, et al. (1994) Wnt-3a regulates somite and tailbud formation in the mouse embryo. Genes Dev 8: 174-189.

65. Nowotschin S, Ferrer-Vaquer A, Concepcion D, Papaioannou VE, Hadjantonakis AK (2012) Interaction of Wnt3a, Msgn1 and Tbx6 in neural versus paraxial mesoderm lineage commitment and paraxial mesoderm differentiation in the mouse embryo. Dev Biol 367: 1-14.

66. Yoshikawa Y, Fujimori T, McMahon AP, Takada S (1997) Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm development in the mouse. Dev Biol 183: 234-242.

67. Aulehla A, Wehrle C, Brand-Saberi B, Kemler R, Gossler A, et al. (2003) Wnt3a Plays a Major Role in the Segmentation Clock Controlling Somitogenesis. Developmental Cell 4: 395-406.

68. Showell C, Binder O, Conlon FL (2004) T-box genes in early embryogenesis. Dev Dyn 229: 201-218.

69. Gluecksohn-Schoenheimer S (1944) The Development of Normal and Homozygous Brachy (T/T) Mouse Embryos in the Extraembryonic Coelom of the Chick. Proc Natl Acad Sci U S A 30: 134-140.

70. Martin BL, Kimelman D (2010) Brachyury establishes the embryonic mesodermal progenitor niche. Genes Dev 24: 2778-2783.

71. Martin BL, Kimelman D (2012) Canonical Wnt Signaling Dynamically Controls Multiple Stem Cell Fate Decisions during Vertebrate Body Formation. Developmental Cell 22: 223-232.

72. Takemoto T, Uchikawa M, Yoshida M, Bell DM, Lovell-Badge R, et al. (2011) Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. Nature 470: 394-398.

73. Hytönen MK, Grall A, Hédan B, Dréano S, Seguin SJ, et al. (2009) Ancestral T-Box Mutation Is Present in Many, but Not All, Short-Tailed Dog Breeds. Journal of Heredity 100: 236-240.

74. Deforest ME, Basrur PK (1979) Malformations and the Manx syndrome in cats. Can Vet J 20: 304-314.

Araucana^Rp
S S
E
Control
free vertebrae
S S 1 2 3 4 5 6 7 8 9
10 11
pygostyle
D
Araucana^Rp
C
Control
free vertebrae
pygostyle
B
A

Figure 1: Adult and embryonic skeletal analysis.

(A) Adult Araucana^{Rp} male and female birds shown in a composite image (courtesy of Fritz Ludwig). Note the characteristic rounded rump, lacking tail structures. Skeletons of control (B) and Araucana^{Rp} (C) birds (courtesy of the ACA). The free vertebrae and pygostyle are missing in the Araucana^{Rp} skeleton (arrow). (D, E) E18 embryo stained with Alcian Blue in Araucana^{Rp} (D) and control (E) embryos. Arrowheads indicate lateral processes. Vertebral elements are numbers from the first free vertebrae (1-5). The more posterior vertebral elements (6-11) fuse to form the mature pygostyle after hatching.

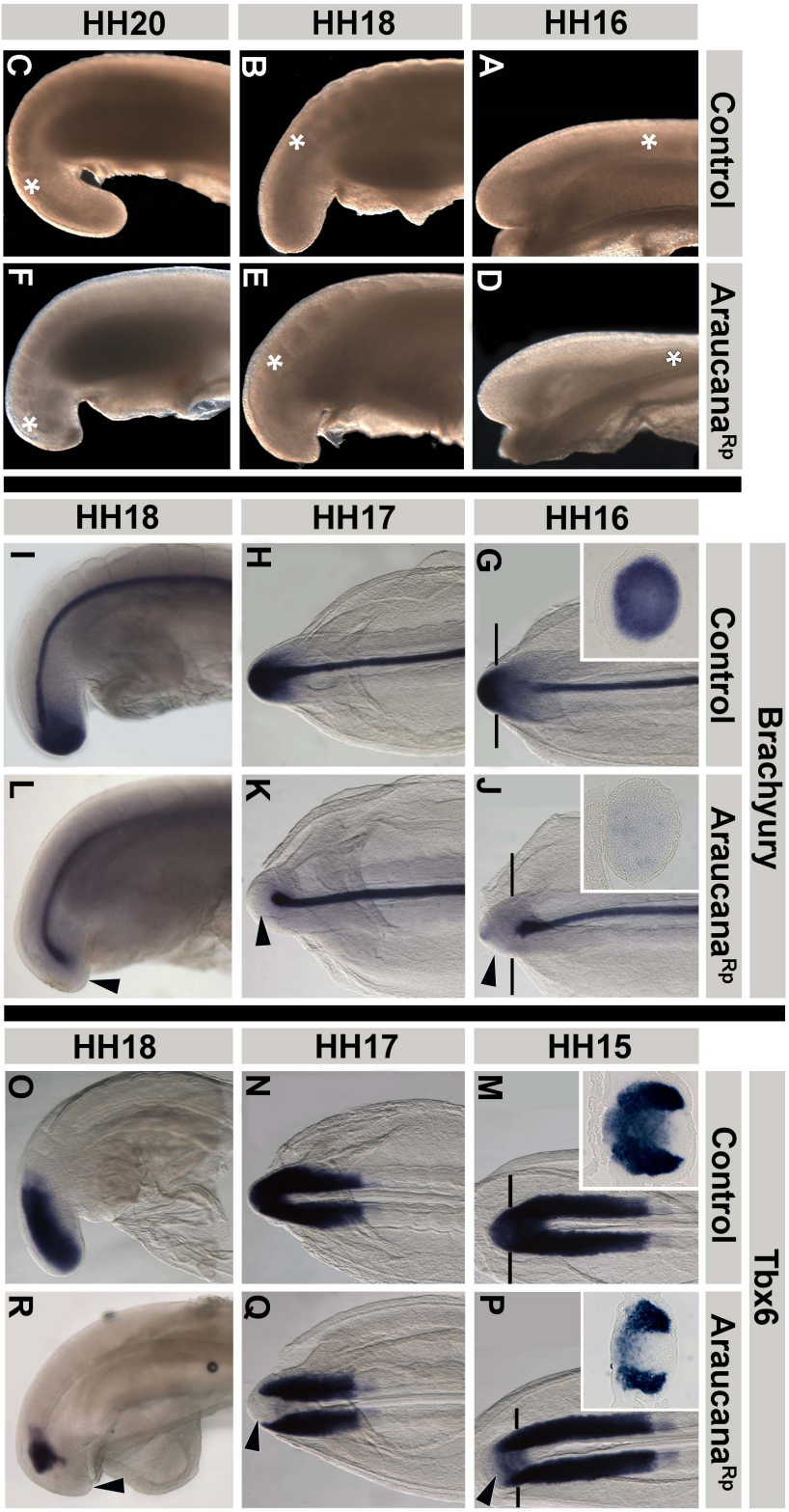

HH16
HH18
HH20
Control
Araucana^Rp
A
B
C
D
E
F
Brachyury
HH16
HH17
HH18
Control
Araucana^Rp
G
H
I
J
K
L
Tbx6
HH15
HH17
HH18
Control
Araucana^Rp
M
N
O
P
Q
R

Figure 2: Araucana^{Rp} embryo tailbuds are truncated and down regulate *Brachyury* and *Tbx6*.

Whole mount tails in lateral view, with posterior at the bottom, dorsal to the left (A-F). (A, D) At HH16, the angle of tail curvature is narrower in controls (A) compared to Araucana^{Rp} embryos (D-arrowed). (B, E) By HH18, the reduced length and pointed shape of the Araucana^{Rp} tail is dramatic (E). (C, F) The control tail has curved ventrally by HH20 (C), whereas the Araucana^{Rp} tail has failed to extend (F). Somite formation in Araucana^{Rp} is near to the tip of the tail (F-asterisk). Expression patterns of *Brachyury* (G-L) and *Tbx6* (M-R) during tailbud development. ISH expression of *Brachyury* in control (G-I) and Araucana^{Rp} (J-L) from HH16-18. Inset in G and J are transverse sections of respective embryos at level of tailbud. Note loss of expression in tailbud mesenchyme in Araucana^{Rp} (arrows) versus controls. ISH expression of *Tbx6* in control (M-O) and Araucana^{Rp} (P-R) at HH15, 17 and 18. Inset in M and P are transverse sections of respective embryos at level of tailbud. Note down regulation of *Tbx6* expression in Araucana^{Rp} (arrows) compared to controls.

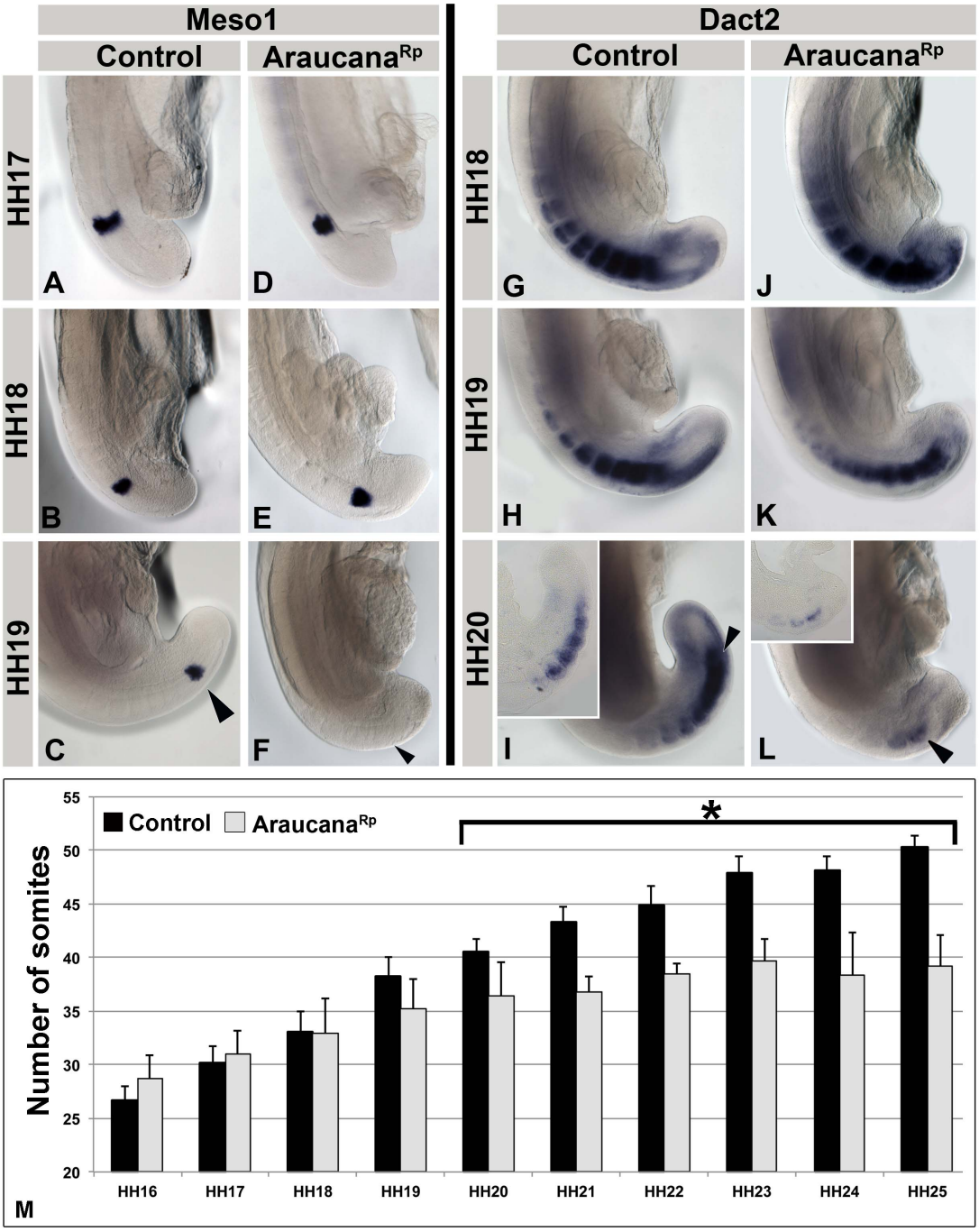

Meso1
Control
Araucana^Rp
Dact2
Control
Araucana^Rp
HH17
HH18
HH19
HH18
HH19
HH20
A
B
C
D
E
F
G
H
I
J
K
L
M
Control
Araucana^Rp
Number of somites
55
50
45
40
35
30
25
20
*
HH16
HH17
HH18
HH19
HH20
HH21
HH22
HH23
HH24
HH25

Figure 3. **Araucana^{Rp} form fewer somites**.

(A-F) *Meso1* expression in control (A-C) and Araucana^{Rp} (D-F) embryos from HH17-19. Note down regulation of *Meso1* in HH19 Araucana^{Rp} (F-arrow) compared to controls (C-arrow). (G-L) *Dact2* expression in control (G-I) and Araucana (J-L) from HH18-20. Inset in I and L are sagittal sections through the paraxial mesoderm from respective embryo. Note down regulation of *Dact2* at HH20 in Araucana^{Rp} (L-arrow) compared to control (I-arrow). Anterior up in whole mount and section insets. (M) Bar graph showing number of somites compared to embryonic(HH16-25). Araucana^{Rp} have significantly fewer somites beginning at HH20 (T-test, asterisk marks $p<0.01$) than controls. Control-Black, Araucana^{Rp}-Grey.

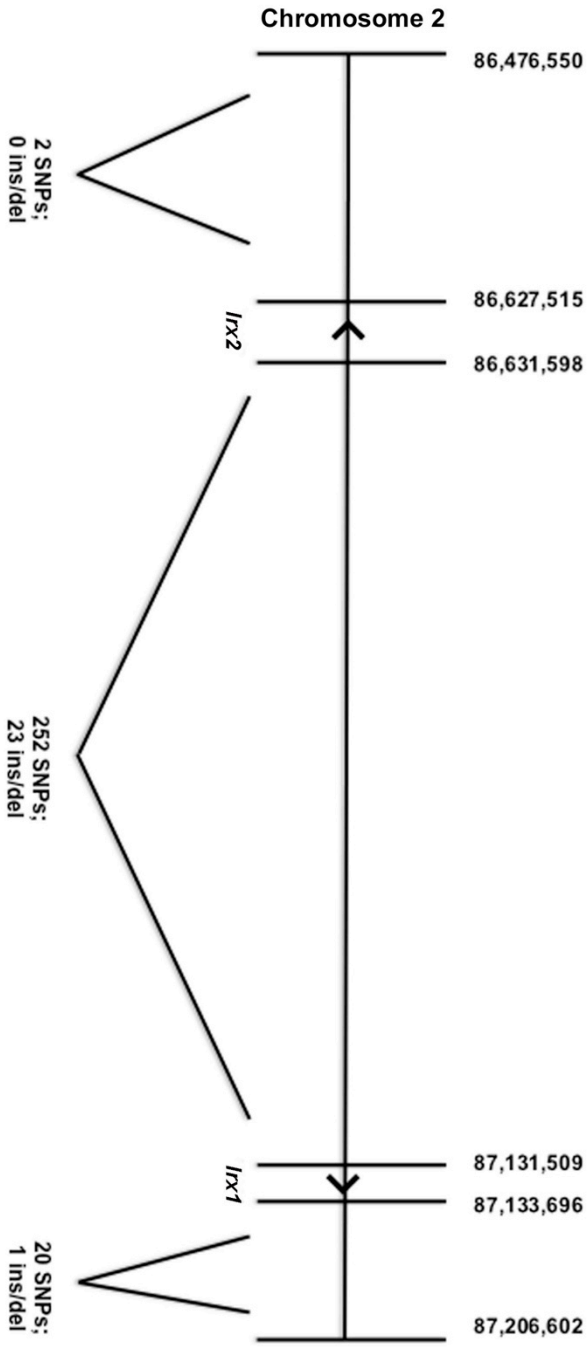

Chromosome 2
86,476,550
2 SNPs;
0 ins/del
86,627,515
Irx2
86,631,598
252 SNPs;
23 ins/del
87,131,509
Irx1
87,133,696
20 SNPs;
1 ins/del
87,206,602

Figure 4: Location of small variants within the 0.74Mb critical region.

Positional map of the 0.74Mb critical region located on *Gallus gallus* chromosome 2. Location of *Irx1* and *Irx2* is indicated, with numbers of small variants (SNPs, insertions and deletions) and their general position indicated. No small variants were found within *Irx1* or *Irx2*.

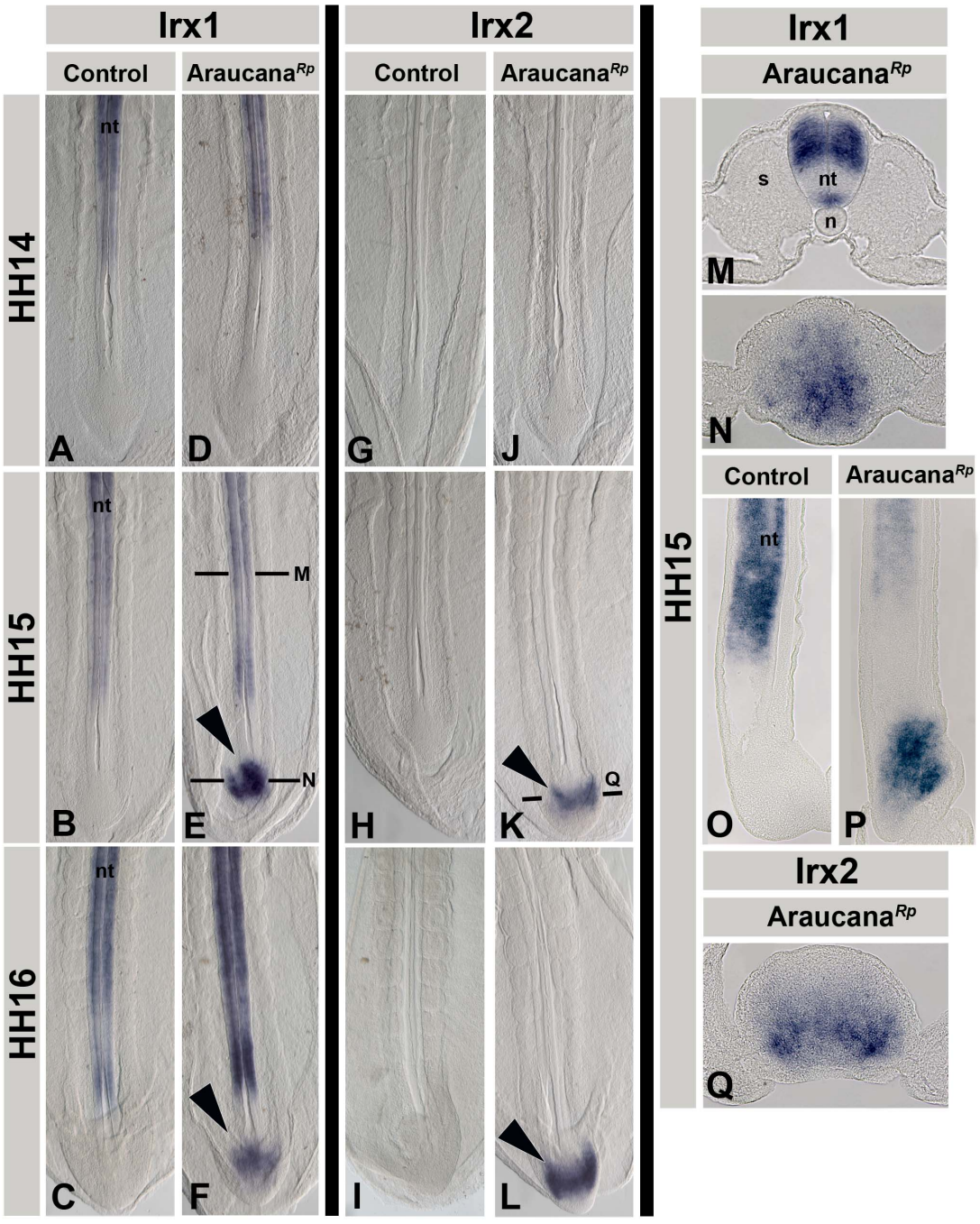

Irx1
Irx2
Irx1
Control
Araucana^Rp
Control
Araucana^Rp
Araucana^Rp
HH14
HH15
HH16
HH15
nt
A
D
G
J
s
nt
n
M
N
B
E
H
K
M
N
l
ql
Control
Araucana^Rp
O
P
C
F
I
L
Irx2
Araucana^Rp
Q

Figure 5: *Irx1* and *Irx2* are misexpressed in Araucana^{Rp} tailbuds.

(A-F) *Irx1* expression pattern. (A-C) Control embryos express *Irx1* in the neural tube. *Irx1* expression in Araucana^{Rp} matches controls at HH14 (D, A), but is misexpressed in the tailbud at HH15 (E-arrow). Normal expression at level of somites is within neural tube (transverse section, M). Misexpression in Araucana^{Rp} can be seen at the level of the chordoneural hinge (transverse section-N and sagittal section-P) compared to expression in HH15 control (sagittal section O). Misexpression is maintained through HH16 in Araucana^{Rp} (F-arrow) as compared to control (C). (G-L) *Irx2* expression pattern. (G-I) Control embryos do not express *Irx2* in the tailbud. No difference in expression between Araucana^{Rp} and controls seen at HH14 (J, G). *Irx2* misexpression in HH15 (K-arrow) and HH16 (L-arrow) Araucana^{Rp}. Transverse section of *Irx2* (Q) shows expression similar to *Irx1* at level of chordoneural hinge. A-L - anterior to top. B,C,K - dorsal to top. M,N,Q - dorsal to top. O,P - dorsal to left, anterior to top. nt-neural tube, cnh-chordoneural hinge, s-somite, n-notochord.

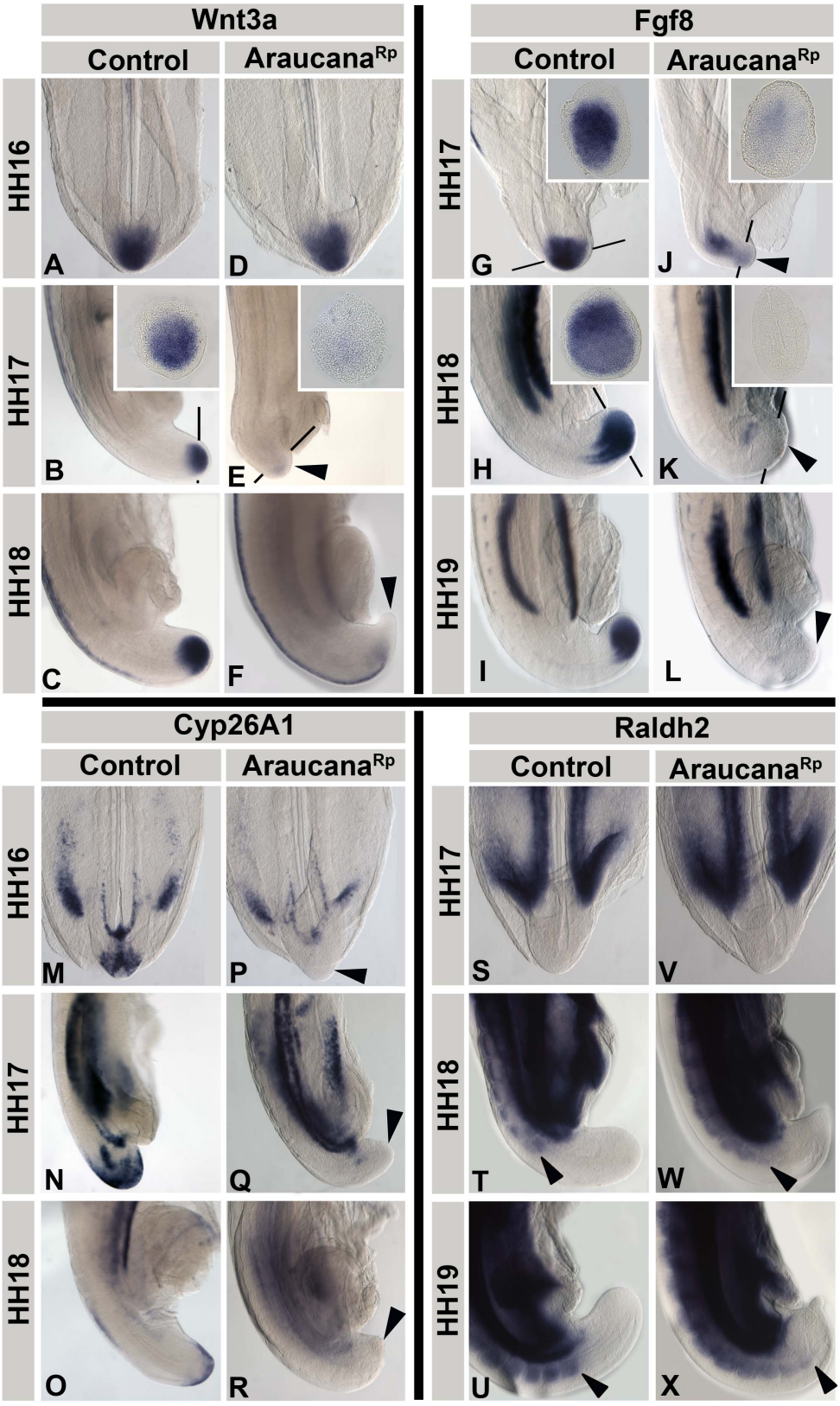

Wnt3a
Control
Araucana^Rp
Fgf8
Control
Araucana^Rp
HH16
HH17
HH18
HH19
A
B
C
D
E
F
G
H
I
J
K
L
Cyp26A1
Control
Araucana^Rp
Raldh2
Control
Araucana^Rp
HH16
HH17
HH18
HH19
M
N
O
P
Q
R
S
T
U
V
W
X

Figure 6. ISH expression pattern of *Wnt3a*, *Fgf8*, *Cyp26A1* and *Raldh2* during tail development.
(A-F) *Wnt3a* expression in control (A-C) and Araucana^{Rp} (D-F) embryos from HH16-18. Inset in B and E are transverse sections at level of tailbud from respective embryo. Note lack of expression in tailbud mesenchyme in Araucana^{Rp} beginning at HH17 (arrows). (G-L) *Fgf8* expression in control (G-I) and Araucana^{Rp} (J-L) embryos from HH17-19. Inset in G, H and J, K are transverse sections at level of tailbud. Note the down regulation of expression in tailbud beginning at HH17 in Araucana^{Rp} (arrows). (M-R) *Cyp26A1* expression in control (M-O) and Araucana^{Rp} (P-R) embryos from HH16-18. Note the lack of expression in Araucana^{Rp} tailbud beginning at HH16 (arrows). (S-X) *Raldh2* expression in control (S-U) and Araucana^{Rp} (V-X) embryos from HH17-19. Note in Araucana^{Rp} the truncated tail at HH19 coupled with the close expression of *Raldh2* in the formed somites (compare U and X). Posterior most expression of *Raldh2* is marked with arrow. Anterior is top in whole mount images. Dorsal is top in transverse section insets. Sections were taken at approximate level of black bars.

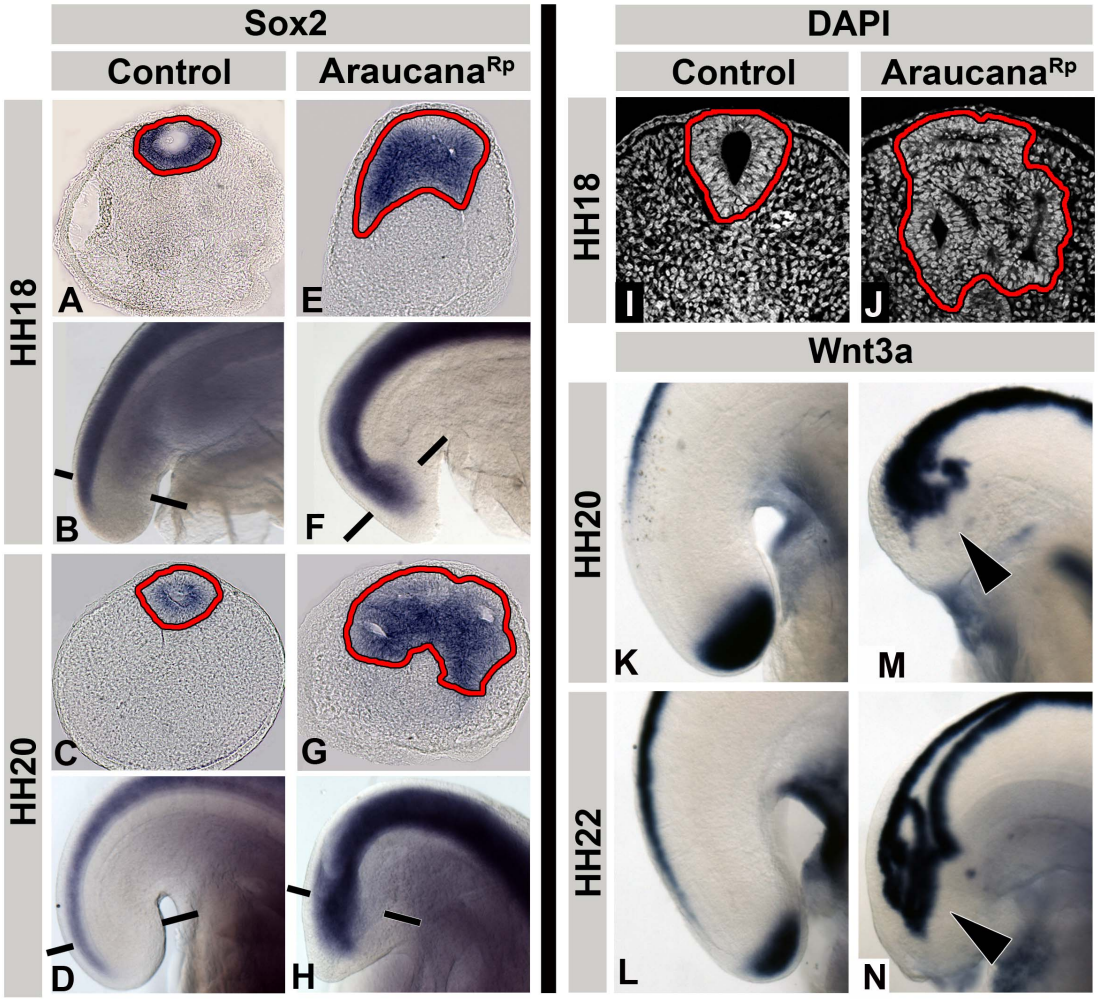

Sox2
Control
Araucana^Rp
HH18
HH20
A
B
C
D
E
F
G
H
DAPI
Control
Araucana^Rp
HH18
I
J
Wnt3a
HH20
HH22
K
L
M
N

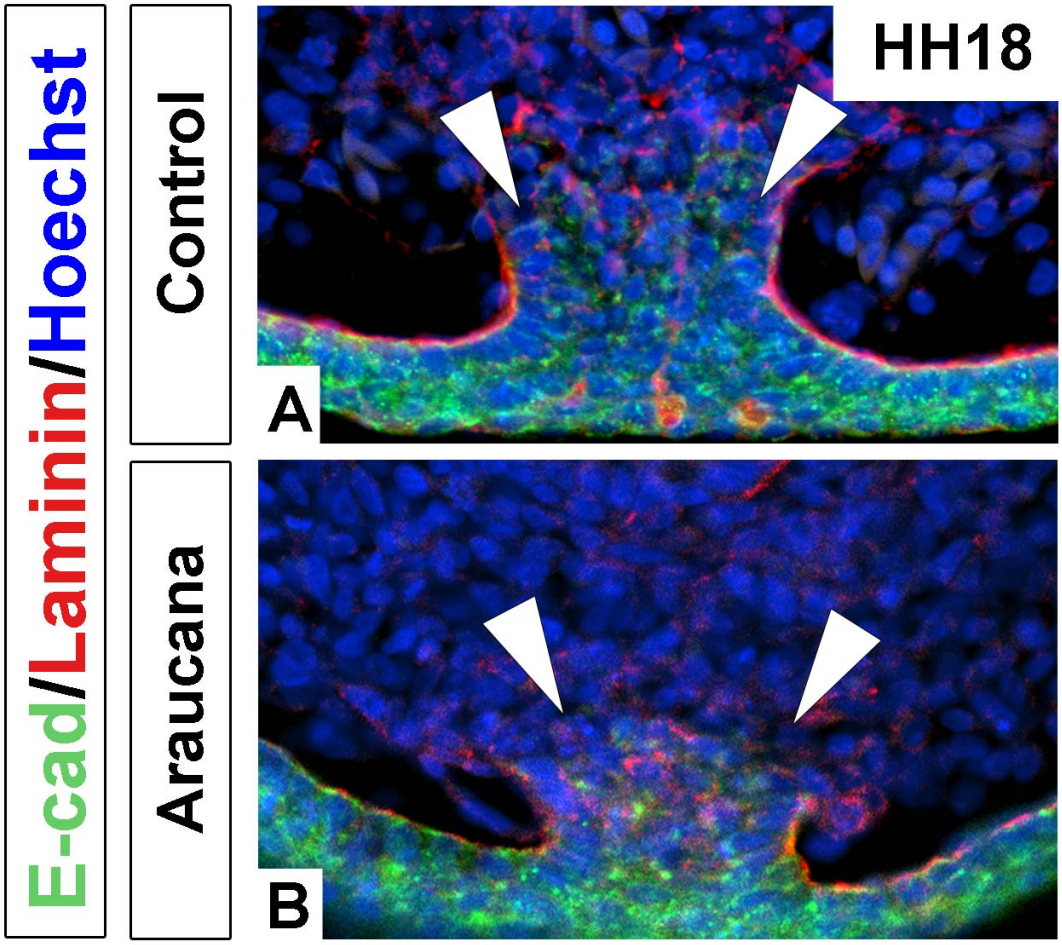
Figure 7: Ectopic neural tubes in Araucana^{Rp} embryo tailbud.

(A-H) Transverse sections and whole mount images showing ISH expression of *Sox2* in control (A-D) and Araucana^{Rp} (E-H) embryos at HH18-20. Note the presence of ectopic neural tissue and lumens in Araucana^{Rp} (highlighted in yellow) (E, G) compared to single neural tube and lumen in controls (A, C). DAPI stained transverse sections of control (I) and Araucana^{Rp} (J) embryo tailbuds. (K-N) Whole mount images of *Wnt3a* expression in control (K, L) and Araucana^{Rp} (M, N). Arrows denote ectopic expression of *Wnt3a* in ectopic neural tubes of Araucana (K, L). Neural tube highlighted. Anterior to top in whole mount images. Dorsal to top in transverse sections. Sections taken at approximate level of black bars.

Edu / TO-PRO 3 Iodide
TUNEL / Hoechst
Araucana^Rp
Control
HH15
HH16
HH17
HH19
A
B
C
D
E
F
G
H
J
K
L
M
N
O
P
Q
I
EdU labeled cells
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
HH14
HH15
HH16
HH17-18
HH19-20
Control
Araucana^Rp
R
TUNEL labeled cells
0
100
200
300
400
500
600
700
HH14
HH15
HH16
HH17-18
HH19-20
Control
Araucana^Rp
*

Figure 8: Role of proliferation and apoptosis in the Araucana^{Rp} tailbud.

(A-H) EdU labeling proliferating cells in sagittal sections from HH15-19 of control (A-D) and Araucana^{Rp} (E-H) embryo tailbuds. Green labels proliferating cells, red is TO-PRO 3 Iodide labeled nuclei. Arrows denote decreased areas of proliferation within the tailbud of Araucana^{Rp} embryos compared to controls. (I) Quantification of the ratio of EdU positive cells to total cells for control and Araucana^{Rp} embryos. (J-Q) TUNEL labeling apoptotic cells in sagittal sections from HH15-19 of control (J-M) and Araucana^{Rp} (N-Q) embryo tailbuds. Red is apoptotic cells, blue is Hoechst labeled nuclei. Arrows denote increased areas of apoptosis in Araucana^{Rp} compared to controls. (R) Quantification of the numbers of TUNEL positive cells in control and Araucana^{Rp}. Araucana^{Rp} significantly more TUNEL positive cells beginning at HH17 (T-test, asterisk marks $p<0.05$) than controls. The same embryo (two different sections) was used for both TUNEL and EdU labeling. Anterior is top, dorsal is left in all sections.


E-cad/Laminin/Hoechst
Control
Araucana
HH18
A
B

Supplemental Figure 1: E-cadherin and laminin staining of the ventral ectodermal ridge.

(A-B) Transverse sections of the tailbud with immunofluorescent labeling of E-cadherin (green), Laminin (red), and Hoechst (blue) in control (A) and Araucana^{Rp} (B) HH18 embryos. Arrows indicate breakdown of basal membrane and ingressing cells through VER in both Araucana and controls. Dorsal to top in sections.

Supplemental Table 1. **Total number of reads, bases, coverage, SNPs, and insertions/deletions within the sequenced region for each bird genotyped.**

Sample	Reads	Bases (kbp)	Coverage	SNPs	Indels
Homo-Rumpless (32)	246,487,326	23,772,362	22.64	723	188
Homo-Rumpless (33)	339,674,102	36,474,356	34.74	750	181
Homo-Rumpless (38)	350,207,902	33,513,411	31,92	751	176
Hetero-Rumpless (39)	235,911,062	22,863,519	21.77	1018	51
Hetero-Rumpless (49)	348,219,696	33,805,751	32.20	466	102
Homo-Tailed (59)	243,759,626	23,629,533	22.50	1434	116
Average	294,043,286	29,009,822	27.63	857	136

Supplemental Table 2. **. Number of unique Araucana small variants found in known highly conserved noncoding regions.** Enhancers were reported in *X. tropicalis* (xenTro2) and coordinates were mapped to the *G. gallus* (Gal/gal4) genome using LiftOver. Enhancer activity reflects what has been previously published [45].

(xenTro2) scaffold_131:	**Enhancer activity**	**Gal/gal4 Coordinates**	**# Small variants found**
752091-754652	-	chr2:86810711-86818277	9 SNPs
764390-767087	-	chr2:86836712-86840304	7 SNPs
944890-945898	-	chr2:86999326-87000129	4 SNPs
979354-980439	-	chr2:87042491-87068516	13 SNPs 2 INSERTIONS

CHAPTER FOUR

DISCUSSION

My research objectives were to identify candidate genes associated with the rumpless phenotype, and to elucidate pathways leading to the rumpless phenotype. I have reported the identification of the chromosome region that is associated with the rumpless phenotype and identified two candidate gain-of-function genes, *IRX1* and *IRX2*. In addition, a number of pathways affecting somitogenesis in Araucana have been identified. These include pathways affecting the maintenance of the tailbud progenitor population and the determination of cell fate of the cells of the tail. This study provides both a novel model of axis truncation defects and an understanding of the mechanisms involved in maintaining the caudal progenitor population.

Analysis of genes in the rumpless haplotype

Using a genome wide association study, I identified a 2.14 Mb haplotype block on chromosome 2 associated to the rumpless phenotype. The haplotype block contained five genes; *CLPTM1L, TERT, IRX1, IRX2, IRX4.* None of these was previously identified as potential upstream effector controlling tail development or in axis truncation defects. Additional analysis of three of the tailed birds found partial blocks of heterozygosity for the rumpless haplotype, delimiting a critical interval of 0.74 Mb. Importantly, this region contained only two genes, *IRX1* and *IRX2*. In situ hybridization analysis of *CLPTM1L,*

TERT and *IRX4* showed no changes in the expected pattern. However, as I have shown, both *IRX1* and *IRX2* exhibit a gain-of-function misexpression pattern in the caudal progenitor population. Investigating the potential role of this mechanism in human caudal malformations, such as caudal regression syndrome, is a long-term goal resulting from this study. Furthermore, this study is significant in that it demonstrates the utility of high throughput mapping using SNPs, and demonstrates that the causes of genetic traits within existing chicken breeds can now be efficiently identified.

iroquois 1 and iroquois 2 homologues

Iroquois 1 (*IRX1*), 2 (*IRX2*) and 4 (*IRX4*) are part of the *Irx* A cluster, with *IRX3, 5* and *6* making up the *IRX* B cluster. *Drosophila* has only one cluster, consisting of *araucan, caupolican* and *mirror*. This suggests a duplication event occurred at some point during evolution between the *Drosophila* and avian radiations. Further evidence for this duplication comes from protein similarity for IRX1 and 3, IRX2 and 5, and IRX4 and 6, indicating they are most likely paralogs [1]. Both *IRX* clusters have been identified in chickens, however, annotation is not fully complete, as cross species comparisons have found that *IRX5* in chicken is more closely related to *IRX6*, indicating chicken *IRX5* annotation should be changed to *IRX6* [2]. *IRX1* and *IRX2* have been identified correctly.

The iroquois genes encode transcription factors with a highly conserved three amino acid loop extension (TALE) in addition to a 13 amino acid Iro box domain. Studies using *Drosophila* have shown iroquois genes are involved in neural patterning through the

regulation of other proneural genes [3,4]. Similarly, work in *Xenopus* embryos has shown that *IRX1* is necessary for the formation of the neural plate through down regulation of *BMP4* [5,6]. Interestingly, the down regulation of *BMP4* occurs through a mechanism involving both *IRX1* and *WNT*, as canonical *WNT* signaling is required for expression of *IRX1* in *Xenopus* [5,6]. This suggests that *IRX1* acts as a downstream mediator between *WNT* and *BMP pathways*. Further evidence for the role of *IRX1* in specification of neural fate comes from experiments in *Xenopus* where injection of *IRX1* was able to promote a neural fate, as seen by induction of *SOX2* expression [6]. This suggests that IRX1 is sufficient to promote neural fate. Loss-of-function experiments revealed that *IRX1* is also required in zebrafish for proper neural formation [7].

Similarly, loss of the entire *IRX* B cluster in mouse leads to malformations in forebrain and spinal cord [8]. In chicken embryos, *IRX1* is expressed in the midbrain and hindbrain, as well as the spinal cord. Expression of *IRX2* is limited primarily to the hindbrain, where expression analysis suggests that it is most likely involved in the formation of the cerebellum and subdivision of the midbrain and hindbrain. In addition, all of six *IRX* genes are expressed in the forming digits of mouse and chicken embryos, suggesting a role in digit identity and patterning [2]. These data suggest that the *IRX* genes play a specific role in regionalization and specification of neural domains.

Upstream regulatory control of *IRX1* and *IRX2* has not been fully elucidated. During *Drosophila* eye development iroquois genes act as dorsal selectors, and are regulated by

hedgehog*, wingless*, and the JAK/STAT pathway [9,10]. In vertebrates, the GSX proteins have been shown to repress iroquois gene expression in *Xenopus*, although no direct binding to the iroquois regulatory region was demonstrated [11]. The iroquois genes act on downstream genes by binding a unique motif, leading to transcriptional repression [12]. Work in *Xenopus* identified Xiro1 as a repressor of *BMP4*, most likely through direct binding to the promoter site [5]. Based on my work, several predictions could be made regarding upstream regulation of *IRX1* and *IRX2*. If misexpression of both genes were due to a mutation in a regulatory region such as an enhancer, this would suggest the presence of a new activator, or the loss of a repressor.

This is of great interest, as identification of novel activators/repressors is important to future studies where manipulation of expression specifically within the chordoneural hinge is required. Identification of a mutation within a regulatory site would also allow for the identification of regulatory transcription factor(s) that normally bind the site during upstream regulation of *IRX1*/*IRX2*. Furthermore, such regions have potential therapeutic benefits. They could be targeted to drive expression of proneural genes, for example, to replace lost neurons in degenerative brain diseases.

Neither *IRX1* nor *IRX2* is normally expressed within the progenitor population during axis elongation. The failure to identify any small variants or mutations that segregate with the rumpless phenotype within the coding region of *IRX1* or *IRX2*, along with the observation that there does not appear to be any change in patterning of the brain, suggest

that both proteins are functioning correctly in these regions, further supporting the case for regionally specific enhancer sites, such as those discovered in *Sox2* [13]. In well-studied models of axis truncation in mice, such as the vestigial tail mouse or the T-brachyury mouse, the truncation results from a loss-of-function mutation, such as the WNT3A hypomorphic mutation of the vestigial tail mouse, and Brachyury loss-of-function mutation in the T mouse [14,15]. The observed misexpression of *IRX1* and *IRX2* in the tailbud of rumpless Araucana suggests that rumplessness is caused by a gain-of-function mutation, which is the first incidence reported of a spontaneous gain-of-function mutant causing truncation of the caudal axis.

It remains to be determined whether combined or separate *IRX1* or *IRX2* expression is sufficient to cause the rumpless phenotype. Furthermore, it is unclear what effect misexpression of *IRX1/IRX2* has on stages outside of the range seen in Araucana. What effect would misexpression of either *IRX1* or *IRX2* have on the tail progenitor population at various times in development? In addition, it is unclear if there is a dose dependent effect from the misexpression of *IRX1/IRX2*. Considering the evidence that the rumpless trait in Araucana may have variable expressivity, it could be predicted that a dose dependent effect would be seen with variable levels of ectopic expression of *IRX1* or *IRX2* [16,17].

The chordoneural hinge has previously been defined in terms of its general location and within a large region of overlapping marker gene expression [18,19,20]. Misexpression of

IRX1 and *IRX2* in Araucana^{Rp} embryos occurs within a critical spatiotemporal window at the onset of secondary body formation. I propose that this misexpression occurs specifically in the chordoneural hinge progenitor population, which indicates that the chordoneural hinge is a distinct and transcriptionally unique population from surrounding tissues. Considering the role of IRX1/2 in promotion of neural identity and in neural patterning, the misexpression of both genes in a progenitor population would be predicted to cause a shift towards a neural fate, which is what I have observed.

Bipotential fate choice of the tail progenitor population

This study has not directly shown through experimentation the effect of ectopic expression of *IRX1*/*IRX2* on the progenitor population of the tail, however, predictions can be made based on the timeline of changes in mesoderm and neural markers within the Araucana^{Rp} tailbud. The bipotential fate model suggests that the cells of the tail are a mixed progenitor population that contributes cells to either the neural tube or mesoderm-derived tissues. *TBX6*, *WNT3A*, and brachyury, all have known roles in specifying mesoderm in the cells of the tail [15,21,22,23,24,25,26].

In zebrafish, *WNT* signaling is both necessary and sufficient to drive specification of the progenitor population towards a paraxial mesodermal fate [24]. Similarly, in mice *WNT3A* is required for generation of paraxial mesoderm. Mice mutant for *WNT3A* exhibited both a loss of somitic tissue and ectopic neural tissue [25]. This suggests the identity of caudal cells is determined by local signaling.

Similarly, in mutant *TBX6* mice, ectopic neural tubes form, rather than paraxial mesoderm [22]. TBX6 acts by inhibiting neural specification by binding the SOX2 N1 enhancer [21]. Considering the role of *TBX6* in inhibiting neural cell fate, down regulation of *TBX6* in Araucana^{Rp} permits the N1 enhancer to ectopically up regulate *SOX2* and switch mesoderm to a neural fate. This is supported in *TBX6* knockout mice, where ectopic *SOX2* expression occurs within the paraxial mesoderm [21,22].

Based on this mouse model, I hypothesized that in Araucana^{Rp} embryos would express ectopic *SOX2* in paraxial mesoderm. Unexpectedly, at HH16, down regulation of *TBX6* expression is only observed in the caudal progenitor mesenchyme, with the presomitic mesoderm continuing to express TBX6 normally. However, the presomitic mesoderm is severely reduced or is absent by HH19/20, at approximately the same stage as when the last somites are formed in Araucana^{Rp} embryos. The failure to maintain the presomitic population supports a model in which down regulation of *TBX6* results in a switch of fate toward the neural lineage, which removes progenitor cells from the cell cycle, which in turn results in failure to generate sufficient presomitic mesoderm to support continuing somitogenesis from the remaining progenitor population.

This pattern of regionally specific down regulation, involving the caudal mesenchyme is different in mouse *TBX6* knockouts where TBX6 expression is entirely lost. However, paraxial mesoderm also switches fate, forming ectopic neural tubes [21,22,27]. During secondary body formation, the tail of the mouse *TBX6* knockout does begin to form, but

is truncated by the generation of a mass of undifferentiated mesenchymal cells and ectopic neural tubes [21,22,27]. This suggests that *TBX6* is required for specification or maintenance of mesoderm identity, and also to suppress neural identity, similar to the caudal progenitor mesenchyme population in Araucana^{Rp}.

Based on the finding from the above studies and the observation that all three genes are down regulated within the tail of Araucana^{Rp} I suggest that the remaining progenitor cells within the tail adopt a neural rather than mesodermal fate. Examination of the posterior tail following down regulation of the previously mentioned mesodermal signals revealed ectopic neural tissue expressing *SOX2*. This supports a model in which Araucana^{Rp} progenitor cells have a bipotential fate choice, and a combination of proneural gene misexpression and loss of mesoderm maintenance factors directs cell towards a neural fate. However, this model is currently based on observational data and still requires functional experimental support.

Changes in migration and proliferation

Primary body formation occurs through a combination of processes, including proliferation, ingression through the primitive streak and cell migration in the presomitic mesoderm among others [28]. However, mechanisms directing the origin and migration of cells that contribute to secondary axis have not been fully elucidated. The remaining primitive streak becomes the ventral ectodermal ridge during caudal axis elongation. Cell ingression is required for proper formation of the tail [29]. In Araucana^{Rp}, the basal

lamina of the ventral ectodermal ridge is degraded, theoretically allowing ingression to take place. However, it is unclear if there is an effect on the number of ingressing cells in Araucana^{Rp} embryos, as the breakdown of the basal lamina does not appear identical to controls. Furthermore, the effect of changing the rate or number of cells ingressing on tail elongation remains to be determined.

Early studies on proliferation within the elongating tail revealed a general rate of proliferation with no heavily mitotic population identified in chicken embryos [30,31]. More recent work has found that the chordoneural hinge population is proliferating and self renewing, whereas the more posterior tailbud mesenchyme is not [32,33]. Proliferation in Araucana^{Rp} embryos appears normal until HH17, when the location of proliferating cells between tailed and rumpless embryos diverges. Araucana^{Rp} embryos have very little proliferation within the posterior tailbud population and chordoneural hinge. The decrease in proliferation in this crucial population is further evidence that cells of the progenitor population are switching to a neural fate, undergoing premature differentiation and exiting the cell cycle.

Maintenance of the tailbud progenitor population

The end of somitogenesis normally occurs in chick at approximately HH25, when the remaining unsegmented presomitic mesoderm expresses *RALDH2*, leading to exposure of the tailbud to retinoic acid, coinciding with increased levels of apoptosis [34,35]. The current hypothesis is that apoptosis removes any remaining unsegmented presomitic

mesoderm and progenitor cells through exposure to retinoic acid [31,34,35]. In Araucana^{Rp} embryos there is no change in regional expression of *RALDH2*, rather there is a down regulation of *CYP26A1* in the tailbud at HH16, indicating that premature apoptosis results from a failure to protect the tailbud from retinoic acid, rather than aberrant expression of RALDH2. Interestingly, although the Araucana model does not appear to recapitulate the precise mechanism of axis termination seen during normal tail development, it is similar to axis termination in mice. In mouse embryos, although there is an increase in expression of *RALDH2*, it does not coincide with an increase in retinoic acid activity [34]. However, there is a similar decrease in the protective effect of CYP26A1 at the end of somitogenesis, suggesting that in both mouse and Araucana^{Rp} embryos the onset of axis of termination is induced by exposure of the remaining unsegmented presomitic mesoderm to retinoic acid.

Both *TBX6* and *CYP26A1* expression are down regulated in Araucana^{Rp} embryos at HH16. The relative contributions due to loss of *TBX6* in the progenitor population causing a switch in fate, and exposure to retinoic acid due to the down regulation of *CYP26A1* resulting in premature cessation of somitogenesis, remains to be determined. Further studies will elucidate the downstream effects iroquois misexpression, and determine if iroquois can directly bind and regulate either *TBX6* or *CYP26A1*.

In conclusion, I have identified a novel gain-of-function mutation in Araucana^{Rp} embryos that results in misexpression of the proneural IRX1 and IRX2 genes within the caudal

progenitor population required for proper secondary body formation. This triggers a cascade of events, including altered gene expression patterns, failure of cell fate maintenance, cell fate switches, changes in migration, proliferation and apoptosis, which together result in axial truncation.

The Araucana chicken breed, carrying the rumpless phenotype has proved to be a valuable model in identifying a novel genetic mutation and in understanding axial elongation. This model promises to continue informing our understanding of normal secondary body axis formation and elucidating the genetic and molecular mechanisms required for axial elongation. Moreover, the malformation observed in this breed may be critical to understanding human caudal pathogenesis. Finally, ongoing studies to identify the causative mutation offers the promise of identifying important genetic targets for treating both congenital malformations and manipulating cell fate and proliferation in degenerative brain disorders.

References

1. Gomez-Skarmeta JL, Modolell J (2002) Iroquois genes: genomic organization and function in vertebrate neural development. Curr Opin Genet Dev 12: 403-408.

2. McDonald LA, Gerrelli D, Fok Y, Hurst LD, Tickle C (2010) Comparison of Iroquois gene expression in limbs/fins of vertebrate embryos. J Anat 216: 683-691.

3. Gomez-Skarmeta JL, Diez del Corral R, de la Calle-Mustienes E, Ferre-Marco D, Modolell J (1996) Araucan and caupolican, two members of the novel iroquois complex, encode homeoproteins that control proneural and vein-forming genes. Cell 85: 95-105.

4. Leyns L, Gomez-Skarmeta JL, Dambly-Chaudiere C (1996) iroquois: a prepattern gene that controls the formation of bristles on the thorax of Drosophila. Mech Dev 59: 63-72.

5. Glavic A, Gomez-Skarmeta JL, Mayor R (2001) Xiro-1 controls mesoderm patterning by repressing bmp-4 expression in the Spemann organizer. Dev Dyn 222: 368-376.

6. Gomez-Skarmeta J, de La Calle-Mustienes E, Modolell J (2001) The Wnt-activated Xiro1 gene encodes a repressor that is essential for neural development and down regulates Bmp4. Development 128: 551-560.

7. Itoh M, Kudoh T, Dedekian M, Kim CH, Chitnis AB (2002) A role for iro1 and iro7 in the establishment of an anteroposterior compartment of the ectoderm adjacent to the midbrain-hindbrain boundary. Development 129: 2317-2327.

8. Peters T, Ausmeier K, Dildrop R, Ruther U (2002) The mouse Fused toes (Ft) mutation is the result of a 1.6-Mb deletion including the entire Iroquois B gene cluster. Mamm Genome 13: 186-188.

9. Zeidler MP, Perrimon N, Strutt DI (1999) Polarity determination in the Drosophila eye: a novel role for unpaired and JAK/STAT signaling. Genes Dev 13: 1342-1353.

10. Cavodeassi F, Diez Del Corral R, Campuzano S, Dominguez M (1999) Compartments and organising boundaries in the Drosophila eye: the role of the homeodomain Iroquois proteins. Development 126: 4933-4942.

11. Winterbottom EF, Ramsbottom SA, Isaacs HV (2011) Gsx transcription factors repress Iroquois gene expression. Dev Dyn 240: 1422-1429.

12. Bilioni A, Craig G, Hill C, McNeill H (2005) Iroquois transcription factors recognize a unique motif to mediate transcriptional repression in vivo. Proc Natl Acad Sci U S A 102: 14671-14676.

13. Kondoh H, Takemoto T (2012) Axial stem cells deriving both posterior neural and mesodermal tissues during gastrulation. Curr Opin Genet Dev 22: 374-380.

14. Greco TL, Takada S, Newhouse MM, McMahon JA, McMahon AP, et al. (1996) Analysis of the vestigial tail mutation demonstrates that Wnt-3a gene dosage regulates mouse axial development. Genes & Development 10: 313-324.

15. Yamaguchi TP, Takada S, Yoshikawa Y, Wu N, McMahon AP (1999) T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. Genes & Development 13: 3185-3190.

16. Landauer W, Dunn LC (1925) Two Types of Rumplessness in Domestic Fowls: A Morphological Comparison. Journal of Heredity 16: 153-160.

17. Dunn LC, Landauer W (1934) The genetics of the rumpless fowl with evidence of a case of changing dominance. Journal of Genetics 29: 217-243.

18. Catala M, Teillet M-A, Le Douarin NM (1995) Organization and development of the tail bud analyzed with the quail-chick chimaera system. Mechanisms of Development 51: 51-65.

19. Charrier JB, Teillet MA, Lapointe F, Le Douarin NM (1999) Defining subregions of Hensen's node essential for caudalward movement, midline development and cell survival. Development 126: 4771-4783.

20. Cambray N, Wilson V (2007) Two distinct sources for a population of maturing axial progenitors. Development 134: 2829-2840.

21. Takemoto T, Uchikawa M, Yoshida M, Bell DM, Lovell-Badge R, et al. (2011) Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. Nature 470: 394-398.

22. Chapman DL, Papaioannou VE (1998) Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. Nature 391: 695-697.

23. Martin BL, Kimelman D (2010) Brachyury establishes the embryonic mesodermal progenitor niche. Genes Dev 24: 2778-2783.

24. Martin BL, Kimelman D (2012) Canonical Wnt Signaling Dynamically Controls Multiple Stem Cell Fate Decisions during Vertebrate Body Formation. Developmental Cell 22: 223-232.

25. Takada S, Stark KL, Shea MJ, Vassileva G, McMahon JA, et al. (1994) Wnt-3a regulates somite and tailbud formation in the mouse embryo. Genes Dev 8: 174-189.

26. Yoshikawa Y, Fujimori T, McMahon AP, Takada S (1997) Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm development in the mouse. Dev Biol 183: 234-242.

27. White PH, Farkas DR, McFadden EE, Chapman DL (2003) Defective somite patterning in mouse embryos with reduced levels of Tbx6. Development 130: 1681-1690.

28. Benazeraf B, Francois P, Baker RE, Denans N, Little CD, et al. (2010) A random cell motility gradient downstream of FGF controls elongation of an amniote embryo. Nature 466: 248-252.

29. Ohta S, Suzuki K, Tachibana K, Tanaka H, Yamada G (2007) Cessation of gastrulation is mediated by suppression of epithelial-mesenchymal transition at the ventral ectodermal ridge. Development 134: 4315-4324.

30. Schoenwolf GC (1977) Tail (End) bud contributions to the posterior region of the chick embryo. The Journal of Experimental Zoology 201: 227-245.

31. Mills CL, Bellairs R (1989) Mitosis and cell death in the tail of the chick embryo. Anatomy and Embryology 180: 301-308.

32. Cambray N, Wilson V (2002) Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. Development 129: 4855-4866.

33. McGrew MJ, Sherman A, Lillico SG, Ellard FM, Radcliffe PA, et al. (2008) Localised axial progenitor cell populations in the avian tail bud are not committed to a posterior Hox identity. Development 135: 2289-2299.

34. Tenin G, Wright D, Ferjentsik Z, Bone R, McGrew M, et al. (2010) The chick somitogenesis oscillator is arrested before all paraxial mesoderm is segmented into somites. BMC Developmental Biology 10: 24.

35. Olivera-Martinez I, Harada H, Halley PA, Storey KG (2012) Loss of FGF-Dependent Mesoderm Identity and Rise of Endogenous Retinoid Signalling Determine Cessation of Body Axis Elongation. PLoS Biol 10: e1001415.